

# **Commonsense Consequentialism**

*Wherein Morality Meets Rationality*

DOUGLAS W. PORTMORE  
*Arizona State University*

*For Erin and Fiona*

---

---

# Contents

0. Front Matter	
0.1 Title Page	
0.2 Dedication	
0.3 Table of contents	
0.4 Acknowledgements	
1. Introduction.....	1
1.1 Utilitarianism: The good and the bad	
1.2 The plan for the rest of the book	
1.3 My aims	
1.4 Objective oughts and objective reasons	
1.5 Conventions that I'll follow throughout the book	
2. Consequentialism and Moral Rationalism.....	25
2.1 The too-demanding objection: How moral rationalism leads us to reject utilitarianism	
2.2 The argument against utilitarianism from moral rationalism	
2.3 How moral rationalism compels us to accept consequentialism	
2.4 What is consequentialism?	
2.5 The presumptive case for moral rationalism	
2.6 Some concluding remarks	
3. The Teleological Conception of Practical Reasons.....	60
3.1 Getting clear on what the view is	
3.2 Clearing up some misconceptions about the view	
3.3 Scanlon's putative counterexamples to the view	
3.4 Arguments for the view	
4. Consequentializing Commonsense Morality.....	91
4.1 How to consequentialize	
4.2 The deontic equivalence thesis	
4.3 Beyond the deontic equivalence thesis: How consequentialist theories can do a better job of accounting for our considered moral convictions than even some nonconsequentialist theories can	

4.4	The implications of the deontic equivalence thesis	
4.5	An objection	
5.	Dual-Ranking Act-Consequentialism: Reasons, Morality, and Overridingness.....	128
5.1	Some quick clarifications	
5.2	Moral reasons, overridingness, and agent-centered options	
5.3	Moral reasons, overridingness, and supererogation	
5.4	A meta-criterion of rightness and how it leads us to adopt dual- ranking act-consequentialism	
5.5	Norcross's objection	
5.6	Splawn's objection	
5.7	Violations of the transitivity and independence axioms	
6.	Imperfect Reasons and Rational Options.....	165
6.1	Kagan's objection: Are we sacrificing rational options to get moral options?	
6.2	Imperfect reasons and rational options	
6.3	Securitism	
6.4	Securitism and the basic belief	
6.5	Securitism's suppositions and implications	
7.	Commonsense Consequentialism.....	210
7.1	The best version of act-utilitarianism: commonsense utilitarianism	
7.2	Securitist consequentialism and the argument for it	
7.3	Commonsense consequentialism and how it compares with traditional act-consequentialism	
7.4	What has been shown and what remains to be shown	
8.	Back Matter	
8.1	Glossary.....	249
8.2	Bibliography.....	270
8.3	Index of names [not completed yet]	
8.4	Index of subjects [not completed yet]	

---

---

## Acknowledgements

This has yet to be written. For now, it's just a log of who I'll need to thank:

- For financial support while working on this book, I thank Arizona State University and the Center for Ethics and Public Affairs, The Murphy Institute, Tulane University, and especially the Institute's Director, Richard F. Teichgraber III, and the Center's Coordinator, Margaret M. Keenan, for their help and support while I was in residence at the Center.
- For helpful comments and discussions on early ancestors of this material, I thank Richard Arneson, David Benatar, Noell Birondo, Dan Boisvert, Ben Bradley, Campbell Brown, Michael Byron, Cheshire Calhoun, Andrew Cochrane, Stew Cohen, Peter de Marneffe, Dale Dorsey, Jamie Dreier, Julia Driver, Nir Eyal, Fred Feldman, J. L. A. Garcia, Josh Gert, Josh Glasgow, Pat Greenspan, Matthew Hanser, Caspar Hare, Chris Heathwood, Brad Hooker, Frances Howard-Snyder, Thomas Hurka, Shelly Kagan, Errol Lord, Jennie Louise, Sharon Lloyd, Eric Mack, John Maier, Peter Marchetto, Chris McMahon, Paul McNamara, David McNaughton, G. Shyam Nair, Robert Noggle, Howard Nye, Sven Nyholm, Jonas Olson, Derek Parfit, Betsy Postow, Peter Railton, Ben Sachs, Geoffrey Sayre-McCord, Mark Schroeder, David Shoemaker, Ted Sider, David Sobel, Jussi Suikkanen, Sergio Tenenbaum, Mark Timmons, Valerie Tiberius, Mark van Roojen, Pekka Väyrynen, Jean-Paul Vessel, Steven Wall, Michael J. Zimmerman, numerous anonymous journal referees, students in my Spring 2008 seminar on consequentialism, and audiences at Tulane University, Arizona State University, the 2008 Bellingham Summer Philosophy Conference, the 2008 Rocky Mountain Ethics Congress, the 2006 APA Pacific Meeting, the 2010 Arizona Workshop on Normative Ethics, and the 2000 and 2005 International Society for Utilitarian Studies conferences.
- For helpful comments on some portion of the manuscript itself, I thank Steve Campbell, Erik Carlson, Richard Chappell, Josh Glasgow, Andrew Khoury, G. Shyam Nair, Andrew Schroeder, Russ Shafer-

*Douglas W. Portmore*

Landau, Paul McNamara, Dave Shoemaker, David Sobel, Andy Specht, Toni Rønnow-Rasmussen, Jean-Paul Vessel..., and especially Peter de Marneffe, Robert Neal, Ben Sachs, Travis Timmerman, and ..., who each read and commented on the entire manuscript.

- For help with logic, I thank Errol Lord, G. Shyam Nair, and especially Peter Marchetto.
- For many helpful discussions, I thank commentators on my posts at *PEA Soup* < <http://peasoup.typepad.com/peasoup/> > as well as Richard Chappell and his commentators on his blog, *Philosophy, et cetera* < <http://www.philosophyetc.net/> >.
- For their love and support while writing this book, I thank my wife and daughter, Erin and Fiona Portmore, and my parents: Linda Hills, Pakinee Portmore, and Ralph Portmore.
- For permission to borrow material from some of my previous publications, I thank the editors and publishers of the following articles:
  - “Consequentialism and Moral Rationalism,” provisionally forthcoming in *Oxford Studies in Normative Ethics*.
  - “Imperfect Reasons and Rational Options,” forthcoming in *Noûs*.
  - “Consequentializing,” *Philosophy Compass* 4 (2009): 329-347.
  - “Are Moral Reasons Morally Overriding?” *Ethical Theory and Moral Practice* 11 (2008): 369-388.
  - “Dual-Ranking Act-Consequentialism,” *Philosophical Studies* 138 (2008): 409-427.
  - “Position-Relative Consequentialism, Agent-Centered Options, and Supererogation,” *Ethics* 113 (2003): 303-332.
  - “Can Consequentialism Be Reconciled with Our Common-Sense Moral Intuitions?” *Philosophical Studies* 91 (1998): 1-19.

# CHAPTER 1

---

---

## Introduction

- 1.1 Utilitarianism: The good and the bad
- 1.2 The plan for the rest of the book
- 1.3 My aims
- 1.4 Objective oughts and objective reasons
- 1.5 Conventions that I'll follow throughout the book

### §1.1 Utilitarianism: The good and the bad

LET ME START with a confession. I used to be a hardcore utilitarian. According to utilitarianism, an act is morally permissible if and only if it maximizes aggregate utility.<sup>1</sup> And I held, rather simplistically, that aggregate utility just consists in the net sum of pleasure minus pain. What made me a *hardcore* utilitarian was my unapologetic stance regarding utilitarianism's counterintuitive implications. When someone would ask, "So, do you think that it's permissible to torture an innocent child if this would maximize aggregate utility?", I would respond: "It's not just permissible; it's *obligatory!*" That's the kind of simplistic and unapologetic utilitarian that I was. My views have, I would like to think, matured since then. Eventually, utilitarianism's counterintuitive implications were too much for even me to stomach.

What led to my conversion? It was not, as one might suppose, the sorts of counterintuitive implications that are typically cited in today's philosophy lectures and textbooks. Although it is often cited as one of utilitarianism's most counterintuitive implications, I don't find it utterly absurd to suppose that agents are, other things being equal, permitted to commit one murder (which I'll define as the intentional killing of an innocent person) so as to prevent five murders. After all, the state of affairs

<sup>1</sup> The aggregate utility produced by an act is the sum of all the utility it produces minus the sum of all the disutility it produces, where utility is a measure of whatever it is that enhances a subject's welfare, and disutility is a measure of whatever it is that diminishes a subject's welfare. An act maximizes aggregate utility if and only if there is no available alternative act that would produce more aggregate utility than it does. And note that I use 'utilitarianism' as shorthand for 'maximizing act-utilitarianism'.

in which five murders have been committed seems to be five times as bad as the one in which only one murder has been committed. Nor do I find it utterly absurd to suppose that agents are morally required to do some very demanding things—such as, donate all their disposable income to hunger-relief organizations—if this will save hundreds of children from starving to death. After all, an agent's material wealth is not nearly as important as the lives of hundreds of children.

What soured me on utilitarianism were not these counterintuitive implications, but the following two. First, utilitarianism implies that agents are morally required to commit murder whenever doing so will maximize aggregate utility, and, thus, even when the net gain would be quite small, even as small as one *utile*. And let me just stipulate that a *utile* is the smallest possible unit of utility: equivalent to someone's experiencing the mildest of pleasures for the briefest of moments. Second, utilitarianism implies that agents should sacrifice, not only their disposable income, but even their own lives and the lives of those whom they most love whenever doing so will maximize aggregate utility, and, thus, even when the net gain would be as small as one *utile*. I find these implications *utterly* absurd.

I find them absurd, for I find it incredible to suppose that such miniscule gains in the aggregate utility could be sufficient to make it reasonable to perform such acts. Of course, I acknowledge that agents have some reason to maximize aggregate utility, but it seems that they have other reasons as well. For instance, it seems that I have a reason to preserve my own life apart from whether or not my doing so would maximize aggregate utility. And it seems that I have a reason to refrain from murder apart from whether or not my doing so would maximize aggregate utility. What's more, it seems that these other reasons sometimes decisively oppose the reason that I have to maximize aggregate utility. If, for instance, my choice is between saving my own life (and let me tell you that it's a good one) and saving some stranger's life, and the only thing that speaks in favor of the latter is that this would result in there being one more *utile* overall, then it seems to me that I have not just sufficient, but decisive, reason to save my own life. The reason I have to favor my own life over that of some stranger's outweighs the reason I have to maximize aggregate utility, for a net gain of one *utile* is simply not enough to rationally justify sacrificing my own life. Yet utilitarianism requires that I make such an unreasonable sacrifice.

The fact that utilitarianism sometimes requires agents to act contrary to what they have decisive reason to do leads me to reject utilitarianism, for I believe that morality is rationally authoritative—that is, I believe that



moral requirements are requirements of reason. Moral requirements seem to do more than just tell us how we must act in order to comply with certain rules. What's more, they tell us how we must act if we are to act as we ought to (all things considered). In this respect, morality differs from many other codes of conduct. Take the law, for instance. The law tells us how we must act in order to comply with certain legal rules, but it doesn't necessarily tell us how we ought, or have most reason, to act—at least, not all things considered. Indeed, there are times when we ought, all things considered, to violate the rules of law, as when one needs to get one's injured child to the emergency room as quickly as possible and can safely proceed through some intersection against the red light. But, unlike legal requirements, moral requirements do seem to specify requirements of reason. For this reason, I reject any moral theory, such as utilitarianism, that requires agents to act contrary to the requirements of reason.

Although I reject utilitarianism, there is, I admit, something quite compelling about it. At utilitarianism's core is a certain compelling idea that explains, I believe, why I and others have been led to accept it in spite of its wildly counterintuitive implications. The compelling idea seems to be that an act's deontic status is determined by the agent's reasons for preferring its outcome (where this includes everything that would be the case were the act to be performed) to those of the available alternatives such that it can never be morally wrong for her to act so as to bring about the outcome that she has most reason to want to obtain.<sup>2</sup> Since actions are the means by which agents affect how things go, it's only natural to suppose that how an agent should act depends on how she should want things to go. Thus, if she has most reason to want the world to turn out one way rather than another way, then she has most reason to act so as to ensure that it turns out that way rather than the other way.<sup>3</sup> If, for instance, she ought to prefer the world in which she has performed  $x$  to the one in which she has performed  $y$ , then she has more reason to perform  $x$  than to perform  $y$ . When this teleological conception of reasons is combined with

<sup>2</sup> Initially, this may not seem to be utilitarianism's compelling idea, but see Chapter 2 for why I think that it is.

All acts are either permissible or impermissible. All permissible acts are either optional or obligatory. And all optional acts are either merely permissible or supererogatory. These are the deontic statuses that an act can have. (As we'll see, in Chapter 4, the idea that an act is supererogatory is, for me, solely a deontic notion and does not depend on whether its performance is, given the agent's motives and intentions, praiseworthy.)

<sup>3</sup> Like many philosophers these days, I believe that not only are there reasons to intend to perform certain actions and to believe certain propositions, but also reasons to desire certain outcomes. See, for instance, PARFIT 2008, SCANLON 1998, and SKORUPSKI 2009.

the idea that moral requirements are requirements of reason such that it is never morally wrong for an agent to do that which she has most reason to do, we get the compelling idea that it is never morally wrong for an agent to act so as to bring about the outcome that she has most reason to want to obtain.

So I'm torn. On the one hand, it seems to me that utilitarianism cannot be right insofar as it sometimes requires agents to act in ways that are contrary to reason. On the other hand, I find utilitarianism's core idea—that how we ought to act is determined by how we ought to want things to go—compelling. This book attempts to resolve the tension. My hope is that the tension can be resolved by accepting utilitarianism's teleological conception of reasons while rejecting utilitarianism's view that what each agent has most reason to want is that aggregate utility be maximized.<sup>4</sup> In adopting a more plausible view about what it is that agents ought to want, I hope to end up with a substantive moral theory that embodies utilitarianism's teleological conception of reasons while avoiding nearly all, if not all, of its counterintuitive implications. This substantive moral theory is what I call *commonsense consequentialism*, and it is the view that I'll be defending. Unfortunately, though, there is a lot of ground that I must cover before I will be able to state the view with sufficient clarity and precision. In the meantime, I'll settle for mapping out the course ahead.

## §1.2 The plan for the rest of the book

In the next chapter, Chapter 2, I'll argue against *traditional act-consequentialism*: the view according to which an act is morally permissible if and only if it maximizes the (impersonal) good. Utilitarianism is but one species of this view, but, as I show, the whole genus is susceptible to the same counterargument that I sketched above.<sup>5</sup> That is, all traditional versions of act-consequentialism, and not just utilitarianism, are guilty of requiring agents to act contrary to what they have most reason to do. We should, therefore, reject all traditional versions of act-consequentialism, for, as I'll argue, we should accept *moral rationalism*: the view an act's deontic status is determined by the agent's reasons for and against performing it,

<sup>4</sup> At this point, one may object that although utilitarianism is committed to a certain view about what is good, it is not committed to any view about what agents ought to want. I address this worry in the next chapter.

<sup>5</sup> Utilitarianism (i.e., maximizing act-utilitarianism) is a species of traditional act-consequentialism (cf. EYAL 2008). Accordingly, utilitarianism is the view that holds both that traditional act-consequentialism is true and that an act maximizes the good if and only if it maximizes aggregate utility.

such that she can be morally required to do only what she has most reason to do, all things considered.

Interestingly, although moral rationalism leads us to reject all traditional versions of act-consequentialism, it also compels us to accept act-consequentialism, more generally—the view that an act's *deontic status* is determined by the agent's reasons for and against preferring its outcome to those of the available alternatives.<sup>6</sup> Very briefly, it does so via the following argument:

- 1.1 An act's deontic status is determined by the agent's reasons for and against performing it, such that, if a subject, S, is morally required to perform an act,  $x$ , then S has most reason to perform  $x$ . (Moral Rationalism)
- 1.2 The agent's reasons for and against performing an act are determined by her reasons for and against preferring its outcome to those of the available alternatives, such that, if S has most reason to perform  $x$ , then, of all the outcomes that S could bring about, S has most reason to desire that  $x$ 's outcome obtains. (The Teleological Conception of Practical Reasons)
- 1.3 Therefore, an act's deontic status is determined by the agent's reasons for and against preferring its outcome to those of the available alternatives, such that, if S is morally required to perform  $x$ , then, of all the outcomes that S could bring about, S has most reason to desire that  $x$ 's outcome obtains.<sup>7</sup> (Act-Consequentialism<sup>8</sup>)

In Chapter 3, I explain and defend the teleological conception of practical reasons (i.e., 1.2 above). I begin by explaining why my formulation of the teleological conception of (practical) reasons (hereafter 'TCR') differs from that of some of its critics (e.g., T. M. Scanlon and Elizabeth Anderson) in eschewing talk of both value and desirability. I then proceed to clear up some common and potential misconceptions

<sup>6</sup> Admittedly, this is a rather unusual way of understanding act-consequentialism. I'll explain my reasons for understanding act-consequentialism in this way in §2.4.

<sup>7</sup> This is written so as not to exclude satisficing versions of act-consequentialism, which deny that it is always obligatory to produce the highest ranked outcome.

<sup>8</sup> Traditional act-consequentialism is a species of act-consequentialism. Accordingly, traditional act-consequentialism is the view that holds both that act-consequentialism is true and that an act produces the outcome that the agent has most reason to want to obtain if and only if it maximizes the good.

about the view. I also rebut Scanlon's putative counterexamples to TCR, cases where putatively "many of the reasons bearing on an action concern not the desirability of outcomes but rather the eligibility or ineligibility of various other reasons" (SCANLON 1998, p. 84). I conclude the chapter by arguing that we should accept TCR, because it has the unsurpassed ability to systematize our considered convictions about practical reasons.

In Chapter 4, I try both to motivate and to undertake what's known as the consequentializing project: the project of accommodating, under an act-consequentialist framework, the sorts of intuitive deontic verdicts that are typically thought to be compatible with only nonconsequentialist theories. I argue that, when both 'outcome' and 'act-consequentialism' are broadly construed, it turns out that any remotely plausible nonconsequentialist theory can be consequentialized. The recipe for consequentializing a nonconsequentialist theory is simple: Take whatever considerations that the nonconsequentialist theory holds to be relevant to determining the deontic statuses of actions and insist that those considerations are relevant to determining the proper ranking of outcomes. In this way, the consequentialist can produce a ranking of outcomes that when combined with her criterion of rightness yields the same set of deontic verdicts that the nonconsequentialist theory yields. The motive for consequentializing is, as I explain, to arrive at a moral theory that comports with our considered moral convictions without having to forfeit, as nonconsequentialist theories do, the compelling idea that an act's deontic status is determined by the agent's reasons for preferring its outcome to those of the available alternatives such that it can never be morally wrong for her to act so as to bring about the outcome that she has most reason to want to obtain.

In Chapter 5, I continue with the consequentializing project by defending a rather peculiar version of act-consequentialism, which I call *dual-ranking act-consequentialism*:

**DRAC** S's performing  $x$  is morally permissible if and only if, and because, there is no available act alternative that would produce an outcome that S has both more moral reason and more reason, all things considered, to want to obtain than to want  $x$ 's outcome to obtain.

Unlike more traditional versions of act-consequentialism, DRAC doesn't take the deontic status of an action to be a function of some *evaluative*

*ranking* of outcomes.<sup>9</sup> Rather, it takes the deontic status of an action to be a function of a non-evaluative ranking that is in turn a function of two auxiliary, evaluative rankings: namely, (1) a ranking of outcomes in terms of how much moral reason the agent has to want each of them to obtain and (2) a ranking of outcomes in terms of how much reason, all things considered, the agent has to want each of them to obtain. I argue that DRAC is promising in that it can accommodate certain features of commonsense morality that no single-ranking version of act-consequentialism can accommodate: supererogation, agent-centered options, and the self-other asymmetry. And I spend the bulk of the chapter defending DRAC against the charge that its dual-ranking structure is unmotivated. I argue that DRAC's dual-ranking structure is forced upon the act-consequentialist given that non-moral reasons can successfully *counter* moral reasons and thereby prevent them from giving rise to moral requirements, a claim that I also defend in the chapter.<sup>10</sup>

In Chapter 6, I defend DRAC and its way of accommodating agent-centered options against Shelly Kagan's objection that if we defend the existence of agent-centered options by arguing that one's reasons for promoting one's own interests can successfully counter the moral reasons that one has for doing what's best for others, then what we end up with is a rational requirement to promote one's own interests, whereas what we wanted was both a moral, *and* a rational, option either to promote one's own interests or to do what's best for others. Contrary to Kagan, I argue that we don't end up with a rational requirement to promote one's own interests—at least, we don't if we adopt, as I argue we should, a certain theory of objective rationality, *viz.*, *securitism*. The main idea underlying this theory is that what it is rational for an agent to do at time,  $t_1$ , is a function of the choice-worthiness of the future courses of action that are, as of  $t_0$ , *securable* by her.<sup>11</sup> On this theory, a subject,  $S$ , will have the option to act either altruistically or self-interestedly at  $t_1$  provided that there is both a reasonably choice-worthy future course of action that is, as of  $t_0$ , *securable*

<sup>9</sup> An evaluative ranking of outcomes is a ranking in terms of the agent's reasons (or some subset of her reasons) for preferring each outcome to its alternatives.

<sup>10</sup> To say that the reasons to  $\varphi$  successfully counter the reasons to  $\psi$  is to say that the reasons to  $\varphi$  prevent the reasons to  $\psi$  from being decisive by, say, equaling, outweighing, undermining, or silencing them. Another possibility is that the reasons to  $\varphi$  are incommensurable with the reasons to  $\psi$  such that there is sufficient reason both to  $\varphi$  and to  $\psi$ .

<sup>11</sup> In Chapter 6, I explain what it is for a course of action to be, as of a given time, *securable* by the agent. Very roughly, an act is, as of  $t$ , *securable* by  $S$  if and only if there is some intention such that, if  $S$  were, at  $t$ , to form that intention,  $S$  would perform the action.

by S in which S acts altruistically at  $t_1$  as well as a reasonably choice-worthy future course of action that is, as of  $t_0$ , securable by S in which S acts self-interestedly at  $t_1$ . I argue that this theory of objective rationality not only allows us to meet Kagan's objection, but also allows us to account for what Joseph Raz calls the *basic belief*: the belief that, in most typical choice situations, the relevant reasons do not require performing some particular alternative, but instead permit performing any of numerous alternatives (1999, p. 100).

In Chapter 7 (the concluding chapter), I spell out what I take to be the most plausible version of consequentialism: viz., commonsense consequentialism. I state the argument for it and explain some of the advantages it has over traditional versions of consequentialism. And I end by summarizing what has been shown and explaining what remains to be shown.



Before proceeding to Chapter 2, I will, in the following three sections, address three preliminary matters. In the next section, §1.3, I'll state my aims: that is, what it is that I aim to accomplish in these pages. Then, in §1.4, I'll explain the sense of 'ought' and 'reason' with which I'll be primarily concerned. And, in the last section, §1.5, I'll describe the conventions that I'll be following throughout the book—conventions regarding citations, symbolic notation, the numbering of propositions, etc.

### §1.3 My aims

Ultimately, I'll be arguing for a certain moral theory, which I call commonsense consequentialism. My aim, though, is not to convince anyone that commonsense consequentialism is the correct moral theory. That would be too ambitious, for the arguments that I'll be presenting here are far from conclusive. In fact, I'm not confident that commonsense consequentialism is the correct moral theory myself. I certainly wouldn't wager any significant sum of money on its being the correct moral theory. There are at least two reasons why I and others should withhold their full endorsement of the view even after taking account of all the arguments in this book. First, there will be some question as to whether commonsense consequentialism can, as I hope, avoid the sorts of counterintuitive implications that plague more traditional versions of consequentialism. Let me explain. Commonsense consequentialism holds that the deontic status

of an action is a function of the agent's reasons for preferring its outcome to those of the available alternatives. If, for instance, an agent ought to prefer the outcome in which she has saved her lover to the outcome in which she has instead saved two strangers, then, according to commonsense consequentialism, it can't be that she ought to save the two strangers rather than her lover. But how do we know which outcome the agent ought to prefer?

Unfortunately, very little has been written on the topic, and there is, as far as I know, no fully worked out theory concerning what agents have reason to want and to prefer.<sup>12</sup> I certainly don't have such a theory. I will, then, have nothing more than our pre-theoretical intuitions to appeal to in trying to determine what it is that agents have reason to want and to prefer. I'll argue that these pre-theoretical intuitions support the contention that commonsense consequentialism avoids the sorts of counterintuitive implications that plague more traditional versions of consequentialism. But, admittedly, this is all rather speculative since I don't have anything close to a fully developed theory about what agents have reason to want and to prefer. We should, therefore, withhold our full confidence in the view until these pre-theoretical intuitions have been vindicated by our best theory about what agents have reason to want and to prefer.

Second, since I won't, for lack of space, endeavor to conduct a thorough assessment of how commonsense consequentialism compares to its leading rivals, we can't be confident that commonsense consequentialism is superior to its rivals. What I hope to do instead is to develop the theory as best I can and to present it as part of a coherent picture concerning our commonsense practical views—that is, as part of a coherent picture concerning our commonsense views about the nature and substance of both morality and rationality. Since I don't have space here to make the necessary comparative assessments, I can't claim that commonsense consequentialism does the *best* job of explaining and systematizing our various moral intuitions, nor that it does the *best* job of cohering with our other intuitions concerning practical matters; I can claim only that it does an *excellent* job in each case.

My aim, then, isn't to convince readers that commonsense consequentialism is the correct moral theory, but only to convince them that it is a promising moral theory—one that deserves to be taken seriously

<sup>12</sup> This is, no doubt, due to the influence that Hume has had. Nevertheless, it is an assumption of this book that, contrary to Hume, reason is the master of, not a slave to, desires (see, e.g., SCANLON 1998, chap. 1),

and one that is worthy of further consideration. If I succeed in achieving even this modest aim, I'll be more than satisfied.

#### §1.4 Objective oughts and objective reasons

Throughout this book, I'll be concerned with what agents *objectively* ought to do—that is, with what they ought to do given what the reason-providing facts about their choice situation happen to be, and so irrespective both of what they take those facts to be and of what their evidence suggests that those facts are likely to be. More precisely, an agent objectively ought to perform some particular alternative if and only if it's the *best* alternative. The best alternative is not necessarily the one that would produce the best outcome. For to say that some alternative is the agent's best alternative is just to say that it's the alternative that she has most reason to perform, and, of course, it's an open question whether an agent has most reason to perform the alternative that would produce the best outcome.<sup>13</sup> Nevertheless, for the purposes of illustration, let's suppose that, in some particular instance, what an agent has most reason to do is to maximize the good. In that case, what she objectively ought to do is to perform the alternative that would, in fact, maximize the good, and this is so even if she believes that some other act would maximize the good, even if her evidence suggests that some other act would maximize the good, and even if she reasonably believes that she ought to refrain from maximizing the good.

When I talk about an agent's reasons, I'm talking about her *objective reasons*. To understand what an objective reason is, consider the following example. Suppose that Ronnie and Bradley have both been invited to the same party. At the party, there will be good food, great company, excellent music, and even some dancing. But whereas Ronnie enjoys dancing, Bradley doesn't.<sup>14</sup> And let's suppose that neither Ronnie nor Bradley is aware of (or has any evidence for) the fact that there will be dancing at the party. Nevertheless, it seems that this fact counts as a reason (i.e., an objective reason) for Ronnie, but not for Bradley, to go to the party. As I would put it, Ronnie has an additional reason to go to the party—namely, that there will be dancing there.

<sup>13</sup> For an excellent discussion of this issue, see HURLEY 2009, especially chapter 5.

<sup>14</sup> I borrow this example from SCHROEDER 2007b. See SCHROEDER 2009 for more on the distinction between objective and subjective reasons as well as for more on the distinction between objective and subjective oughts.



Some philosophers would balk at my putting it in this way. Although they would accept that there *is* an additional reason for Ronnie to go to party in virtue of the fact that there will be dancing there, they would deny that Ronnie *has* this reason to go. On their view, we must distinguish between (i) the existence of a reason for S to perform  $x$  and (ii) S's having a reason to perform  $x$ . On their view, S's *having* a reason to perform  $x$  is a matter of S's "possessing" that reason in virtue of bearing some epistemic relation to that reason, such as being aware of it or having sufficient evidence for it.<sup>15</sup> I, by contrast, think that we can appropriately talk of S's having a reason to perform  $x$  even if S is completely unaware (and lacks any evidence for) any fact that counts in favor of S's performing  $x$ . Let me say, though, that even if I'm wrong about this, I can say everything I want to say only using different words. If I'm wrong, then throughout this book I'll need to replace all talk of S's having a reason, more reason, most reason, or decisive reason to  $\varphi$  with talk of there being a reason, more reason, most reason, or decisive reason for S to  $\varphi$ . And, thus, with regard to the objective sense of 'ought', I would need to say 'S objectively ought to perform  $x$  if and only if there *is* most reason for S to perform  $x$ ' as opposed to 'S objectively ought to perform  $x$  if and only if S *has* most reason to perform  $x$ '. But since I think that it's perfectly fine to talk about S's having a reason to do  $x$  even in those instances in which S is unaware of (and lacks any evidence for) any fact that counts in favor of S's performing  $x$ , I will continue to talk and write in this way. If some readers object, they can make the appropriate substitutions throughout the book.

Returning now to 'ought', we can better understand the objective sense of 'ought' by noting that we use 'ought' in this sense both when we wonder whether we ought to act as we believe that we should and when we assess, on the basis of 20/20 hindsight, whether we ought to have acted as we did. To illustrate, suppose that I'm trying to defuse a bomb and that I'm not sure whether it is the red or the green wire that I should cut. That is, I'm not sure which one I must cut in order to deactivate it and which one, if cut, will detonate it. Even if the evidence that's available to me strongly supports my contention that cutting the green wire will deactivate the bomb, I may still wonder whether I ought to do so. Indeed, as I reach out to cut the green wire, I may wonder aloud: "Is this really what I ought

<sup>15</sup> See SCHROEDER 2008c for a critique of this view—cf. LORD 2009. I agree with the proponents of this view that there can be decisive reason for S to  $\varphi$ , and yet, if S lacks sufficient evidence for believing that this is so, then it would be unreasonable to expect her to  $\varphi$  or to criticize her for failing to  $\varphi$ .

to be doing.” Now let’s suppose that after cutting the green wire I wake up in the hospital, for, as it turns out, I needed to have cut the red wire to deactivate it. I may, then, say to myself: “I ought to have cut the red wire.” In both of these two quoted sentences, I would be using ‘ought’ in what I’m calling the objective sense.<sup>16</sup>

We can talk not only about what an agent objectively ought to do, but also about what she is objectively required to do and about what she is objectively permitted (or not permitted) to do.<sup>17</sup> An agent is objectively required to perform an act if and only if she has decisive reason to perform that act. And she is objectively permitted to perform an act if and only if she does not have decisive reason to refrain from performing it.

Thus far, I’ve been talking about a particular objective ‘ought’: the ‘ought’ of objective rationality—that is, the ‘ought’ that concerns what an agent has most reason to do, *all things considered*. There is also, however, an objective *moral* ought. An agent objectively morally ought to perform some particular alternative if and only if it’s the morally best alternative—that is, the one that she has most *moral* reason to perform.<sup>18</sup> Furthermore, an agent is objectively morally required to perform an alternative if and only if she has decisive moral reason to perform that alternative. And she is objectively morally permitted to perform an alternative if and only if she does not have decisive moral reason to refrain from performing it. The reader should assume that, throughout this book, I’m talking either about

<sup>16</sup> This suggests another way of understanding what it is that an agent objectively ought to do: S objectively ought to perform  $x$  if and only if performing  $x$  is what a normatively conscientious person would do if she faced S’s choice of alternatives and had certain knowledge of all the relevant normative and non-normative facts. And we could say, by contrast, that S subjectively ought to perform  $x$  if and only if performing  $x$  is what a normatively conscientious person would do if she were in the exact same situation that S is in, facing S’s choice of alternatives and all the normative and non-normative uncertainty that goes along with being in S’s epistemic position.

<sup>17</sup> Here, I’m using ‘ought to  $\varphi$ ’ in a sense that doesn’t imply a requirement to  $\varphi$ . For instance, I may tell my dinner companion that she ought to try the duck, but, in so doing, I would not be claiming that she is in any sense required to try the duck. This sense of ‘ought’ can be called the *advisory sense* of ‘ought’. Another example of the advisory sense of ‘ought’ comes from FERRY 2009: “[Y]ou really ought to visit your mother on her birthday, but you have got to at least send a card.” See also MCNAMARA 1996b.

<sup>18</sup> In other words, S objectively morally ought to perform  $x$  if and only if performing  $x$  is what a morally conscientious person would do if she faced S’s choice of alternatives and had certain knowledge of all the relevant normative and non-normative facts. By contrast, S subjectively morally ought to perform  $x$  if and only if performing  $x$  is what a morally conscientious person would do if she were in the exact same situation that S is in, facing S’s choice of alternatives and all the normative and non-normative uncertainty that goes along with being in S’s epistemic position.

what agents objectively morally ought to do or about what agents objectively ought to do, all things considered—unless, of course, I explicitly state otherwise.

Although my focus will be on the objective sense of ‘ought’, I don’t claim that this is the only sense of ‘ought’. Nor do I claim that it is the one that directly guides us in our first-person deliberations about what to do. Indeed, it’s not, for we can base our first-person practical deliberations only on what we *think* is the case, not on what *is*, in fact, the case.<sup>19</sup> Furthermore, there is, at most, only a tenuous connection between an agent’s being objectively morally required to perform an act and its being appropriate to advise her to perform it and/or to blame (or otherwise criticize) her for failing to perform it.<sup>20</sup> So I readily admit that the objective ‘ought’ is neither the one that figures most prominently in our first-person practical deliberations nor the one that figures most prominently in our everyday practices of blaming and advising. Nevertheless, objective oughts and reasons are of the utmost importance, for they are more fundamental than subjective oughts and reasons—or so I will argue. But, first, let me explain why the objective ought is not the one that figures most prominently in our first-person practical deliberations or in our everyday practices of blaming and advising. To see this, it will be helpful to consider the following example, which I borrow, with some minor revisions, from Derek Parfit (2008):

*Mine Shafts:* A hundred miners are trapped underground in one of two mine shafts. Floodwaters are rising, and the miners are in danger of drowning. Sally can close one of three floodgates. Depending both on which floodgate she closes and on which shaft the miners are in, the results will be one of the following six possible outcomes, labeled o<sub>1</sub>-o<sub>6</sub>:

		The miners are in...	
		Shaft A	Shaft B
Sally closes...	Gate 1	o <sub>1</sub> : Sally saves 100 lives	o <sub>2</sub> : Sally saves no lives
	Gate 2	o <sub>3</sub> : Sally saves no lives	o <sub>4</sub> : Sally saves 100 lives
	Gate 3	o <sub>5</sub> : Sally saves 90 lives	o <sub>6</sub> : Sally saves 90 lives

Sally doesn’t know which shaft the miners are in, but she knows all the above as well as the fact that her available evidence suggests that there is a 50% chance that the

<sup>19</sup> Derek Parfit (2008) and Andrew Sepielli (2009) both make this point.

<sup>20</sup> There is, however, an essential connection between an agent’s being objectively morally required to perform *x* and its being appropriate to blame her for *freely and knowledgeably* performing *x*—see § 2.5.

miners are in Shaft A as well as a 50% chance that the miners are in Shaft B. As a matter of fact, though, the miners are in Shaft A.

It's clear that Sally objectively ought to close Gate 1, for Sally should, in this instance, do whatever would bring about the best available outcome, and  $o_1$  is the best available outcome given that the miners are in Shaft A. Nevertheless, Sally, being a morally conscientious person, closes Gate 3. She closes Gate 3 even though she knows that closing Gate 3 is not what she objectively ought to do. She knows that there's a 50% chance that she objectively ought to close Gate 1 as well as a 50% chance that she objectively ought to close Gate 2, but absolutely no chance that she objectively ought to close Gate 3.

We may wonder, though, why a morally conscientious person would do what she knows that she objectively ought not to do. The answer lies with the fact that a morally conscientious person will be concerned not only with whether or not her acts are (objectively) morally permissible but also with how morally good/bad her morally permissible/impermissible acts are.<sup>21</sup> Some morally permissible acts are much better than others. For instance, spending the afternoon volunteering for Oxfam is morally better than wasting the afternoon getting drunk, but both may be morally permissible. And some morally impermissible acts are much worse than others. For instance, committing genocide is much worse than telling a little lie, and killing someone with malice aforethought is morally worse than unintentionally killing someone as the result of negligence.

In light of the fact that permissible/impermissible acts come in varying degrees of moral goodness/badness, a morally conscientious person will do what she knows to be objectively wrong when she knows that trying to do what's objectively right risks committing a much graver wrong.<sup>22</sup> Thus, Sally closes Gate 3, knowing that this is objectively wrong. She does so because she knows that were she to try to do what's objectively right by closing either Gate 1 or Gate 2, she risks saving no one. And although saving only 90 miners is just as objectively wrong as saving no one, the latter is a far worse wrong.<sup>23</sup> Thus, being a morally conscientious person,

<sup>21</sup> I borrow this point from BYKVIST 2009.

<sup>22</sup> I use 'right' and 'wrong' as synonyms for 'permissible' and 'impermissible', respectively.

<sup>23</sup> Impermissibility doesn't come in degrees; an act can't be more (or less) impermissible than another—see, for instance, ALEXANDER 2008. By contrast, the moral goodness of an act does come in degrees. Some morally impermissible acts are morally worse than others. To say that one act is morally worse than another is just to say that there is more moral reason to refrain from performing the one than to refrain from performing the other. Alastair Norcross (2005, 2006) has argued that if the rightness (permissibility?) of actions is

Sally saves 90 miners rather than risk saving none. So, as the above example illustrates, the objective 'ought' is not the one that directly guides the first-person practical deliberations of morally conscientious persons.<sup>24</sup>

The above example also illustrates how it can be inappropriate to blame (or otherwise criticize) someone for doing what's objectively wrong as well as how it can be appropriate to blame someone for doing what's objectively right. We would not, for instance, blame Sally for closing Gate 3 even if she did so knowing that this was objectively wrong. But we would blame her for closing Gate 1 even though this is objectively right. We would blame her, because, in closing Gate 1, she would be taking an unacceptable risk—a risk that no morally conscientious person would take. Since we think that it is *subjectively wrong* to take such risks, we would blame her for closing Gate 1.<sup>25</sup> So, clearly, the objective ought is not the one that is most directly relevant to our assessments of blame.<sup>26</sup>

What about advice? Isn't there some essential connection between what a fully informed and well-meaning advisor would advise an agent to do and what that agent objectively ought to do? This is what some philosophers claim, but they are mistaken.<sup>27</sup> For one, whether a fully informed and well-meaning advisor should advise an agent to act as the agent objectively should depends on how the agent would react to the advice. If, for instance, the advisor knows that the agent would do the opposite of whatever she advises, she should, perhaps, advise the agent to do the opposite of what she objectively ought to do. And note that even

determined by the goodness of their outcomes, then, since goodness is gradable, then rightness must be gradable as well. For a reply to Norcross, see LAWLOR 2009a, LAWLOR 2009b, and Chapter 5.

<sup>24</sup> For a subtle and well-developed account of what does guide a morally conscientious person in her first-person practical deliberations, see ZIMMERMAN 2008.

<sup>25</sup> Roughly speaking, an act is subjectively wrong if and only if an agent in the normal frame of mind would be blameworthy for performing it—see GIBBARD 1990, pp. 44-45.

<sup>26</sup> Although the objective 'ought' is not the one that's most salient to our *assessments* of blame, there is an essential connection between the objective ought and blameworthiness—see §2.5.

<sup>27</sup> At times, Derek Parfit (2008) seems to make this mistake, equating what I objectively ought to do with what "any fully informed adviser ought to have told me that I ought to do." Joshua Gert (2004, p. 140) gives an account of the objective rational status of an action in terms of its having consequences that virtually everyone could see as allowing "someone to sincerely advise someone else to do it." Frank Jackson and Robert Pargetter (1986, p. 237) say, "If the *right* moral advice is to say no..., then what else can we conclude but that he [objectively] ought to say no?" And Jennie Louise (2008) claims that it is my objective obligation that I  $\phi$  in circumstances C "if and only if my fully informed and rational counterpart would advise me to  $\phi$  in C."

here the connection is somewhat tenuous as the qualifier ‘perhaps’ is meant to indicate, for whether a fully informed and well-meaning advisor should advise an agent to do  $x$  (even assuming that the advice will be taken) depends on whether advising that agent to do  $x$  is *the advisor’s* best alternative, not on whether  $x$  is *the agent’s* best alternative, and it’s only the latter that determines what the agent objectively ought to do.

To illustrate, suppose that, in *Mine Shafts*, a fully informed and well-meaning advisor named Rick is in a position to advise Sally which gate to close. Further suppose that the evil, but trustworthy, Boris has promised Rick that unless he advises Sally to close Gate 2, he’s going to kill a thousand other innocent people. In this case, it seems that Rick should advise Sally to do what’s objectively wrong, viz., to close Gate 2. So, from the fact that Rick, a fully informed and well-meaning advisor, ought (in every sense) to advise Sally to do  $x$ , it does not follow that Sally objectively ought to do  $x$ .

It may seem, though, that at least this much is true: if Sally objectively ought to do  $x$ , then Rick ought to advise Sally to do  $x$  if his doing so would have the direct effect of causing Sally to do  $x$  and no other direct effects, such as causing Boris to kill a thousand other innocent people.<sup>28</sup> But even this much is debatable, for we might think that whereas what Rick ought to do depends on what’s his best alternative, what Sally ought to do depends on what’s her best alternative. And it may be that, although the best alternative for Rick is to advise Sally to do  $x$ ,  $x$  is not Sally’s best alternative.

To illustrate, consider a different case—one that again involves Rick in the role of advisor and Sally in the role of advisee, but one that doesn’t involve any miners. Suppose that, in this new case, Rick must advise Sally as to whether or not to hit her daughter. Assume that this is the only issue on which Rick can advise Sally and that Sally will hit her daughter if and only if Rick advises her to do so. Further suppose that if Sally doesn’t hit her daughter, she’s going to kill her son. Assume that she *could* refrain both from hitting her daughter and from killing her son, as all she needs to do is to decide now to go straight to the gym and get her aggression out by hitting a punching bag. But since Sally has no intention of going to the

<sup>28</sup> We should allow, though, that her advice may have other indirect effects—i.e., the effects that Sally’s doing  $x$  would have.

gym, it is, nevertheless, true to say that what she *would* do if she refrains from hitting her daughter is kill her son.<sup>29</sup>

Now, Rick's best alternative is to advise Sally to hit her daughter, for the only alternative for him is to refrain from giving her this advice; in which case, Sally will do much worse, killing her son. In fact, no matter what Rick does, Sally is going to hurt one of her two children. Given this, the best alternative for Rick is to advise Sally to hit her daughter—at least, that way no one gets killed. But, unlike Rick, Sally can ensure that she doesn't hurt anyone. All she has to do is to decide now to go the gym and get her aggression out by hitting a punching bag. So Sally's best alternative involves her venting her anger, not on her children, but on some punching bag. Here, then, is a case where the best alternative for Rick is to advise Sally to hit her daughter and yet Sally's hitting her daughter is not her best alternative. We might conclude, therefore, that sometimes an advisor should advise an advisee to do something that the advisee objectively ought not to do (FELDMAN 1986, p. 56).<sup>30</sup>

We've seen, then, that whether an agent objectively ought to do  $x$  is not what's most important in determining whether it's appropriate to blame her for doing  $x$  or even in determining whether one should advise her to do  $x$ . We've also seen that the objective sense of 'ought' is not the one that figures most prominently in our first-person deliberations about what to do. The reason we can't base our first-person practical deliberations on what we objectively ought to do is that we are always to some degree uncertain about what we objectively ought to do.<sup>31</sup> There are two sources of this uncertainty: *non-normative uncertainty* (i.e., uncertainty about matters of non-normative fact) and *normative uncertainty* (i.e., uncertainty about matters of normative fact).<sup>32</sup> Given this uncertainty, it follows that we sometimes *subjectively ought* to do what we know to be objectively impermissible. This follows given the following plausible principle:

<sup>29</sup> I'm using 'could' and 'would' in their ordinary senses. For instance, it seems perfectly natural for someone to say: "I could hit my child, but that's not something that I would ever do."

<sup>30</sup> I'll have much more to say about this sort of case in Chapter 7.

<sup>31</sup> For instructive and subtle accounts of what we ought (in some non-objective sense) to do when we don't know what we ought to do in the objective sense, see SEPIELLI 2008, SEPIELLI 2010, and ZIMMERMAN 2008. And see BYKVIST 2009 and MOLLER 2009 for some poignant criticisms of ZIMMERMAN 2008.

<sup>32</sup> I borrow this point along with this way of formulating the distinction between normative and non-normative uncertainty from SEPIELLI 2008.

**P1** If an agent believes that her  $\varphi$ -ing has more expected *deontic value* than any other alternative, then she subjectively ought to  $\varphi$ .<sup>33</sup>

The deontic value of an act is a measure of how much (objective) reason there is to perform it, and we get the expected deontic value of an action “by multiplying the subjective probability that some practical comparative is true by the [deontic] value of that action if it is true, doing the same for all other practical comparatives, and adding up the results”—*practical comparatives* are anything of the form: the balance of reasons favors *S*’s doing *x* as opposed to *y* (SEPIELLI 2009, pp. 7 & 11).<sup>34</sup>

To illustrate the plausibility of P1, we need only look to examples such as *Mine Shafts*. Sally knows that closing Gate 3 is objectively impermissible. But she also knows that closing Gate 3 is the alternative with the highest expected deontic value, for Sally is, we’ll assume, certain that the balance of reasons favors her closing whichever gate would result in the most miners being saved. Given her certainty with regard to this normative question, we can represent the expected deontic value of each of her alternatives quite simply as a function of the probability that the miners are in a particular shaft multiplied by the number of miners that would be saved by closing a given gate on the assumption that they’re in that shaft, which is what I’ve done in Table 1.1. Since Sally knows that her closing Gate 3 has the highest expected deontic, it is, given P1, subjectively wrong for Sally to do anything but close Gate 3.

**Table 1.1**

<sup>33</sup> If the reader prefers, we could formulate this principle as well as the following principles in terms of what the agent ought to believe given her evidence as opposed to what she actually believes, as I’ve done here. In that case, the reader should just replace ‘reasonably believes’ for ‘believes’ where appropriate. For why I’ve chosen to formulate these principles in terms of the agent’s actual beliefs as opposed to her reasonable beliefs, see H. M. SMITH 2011.

<sup>34</sup> For a fuller account, see SEPIELLI 2009. To avoid confusion, I have substituted Sepielli’s term ‘objective value’ with Zimmerman’s less misleading term ‘deontic value’ (ZIMMERMAN 1996, pp. 14-15). Regardless of which term we use, the notion to which it refers is meant to be a theory-neutral one. An act’s objective/deontic value is just a measure of how much (objective) reason there is for the agent to perform it. Thus, we can hold that an agent subjectively ought to do what has the highest expected deontic value without committing ourselves to any substantive view about what she subjectively ought to do. Also, along with the notion of an act’s deontic value, there’s the notion of an act’s deontic moral value, which is a measure of how much moral reason there is for the agent to perform it. Thus, the moral analogue of P1 is: ‘If an agent believes that her  $\varphi$ -ing has more expected deontic moral value than any other alternative, then she subjectively morally ought to  $\varphi$ ’.



<i>Alternative</i>	<i>Expected Deontic Value</i>
Closing Gate 1	$(.5 \times 100) + (.5 \times 0) = 50$
Closing Gate 2	$(.5 \times 0) + (.5 \times 100) = 50$
Closing Gate 3	$(.5 \times 90) + (.5 \times 90) = 90$

Of course, P1 isn't the only plausible principle concerning what an agent subjectively ought to do. The following principles are also plausible:

- P2** If an agent believes that she is objectively required to  $\varphi$  (where  $\varphi$  is some positive act and not the "act" of refraining from performing some positive act), then she is subjectively required to  $\varphi$ .<sup>35</sup>
- P3** If an agent believes that there is a fair chance that her refraining from  $\varphi$ -ing is objectively wrong and that there is no chance that her  $\varphi$ -ing is objectively wrong, and if her  $\varphi$ -ing would be no more personally costly than her refraining from  $\varphi$ -ing, then she is subjectively required to  $\varphi$ .
- P4** If an agent believes that she is likely to face the choice of whether or not to  $\psi$ , that whether or not she is objectively required to  $\psi$  depends on whether or not  $p$  is true, and that by  $\varphi$ -ing she could come to be in a better epistemic position with respect to  $p$ , then she is, other things being equal, subjectively required to  $\varphi$ .

Of course, this list isn't meant to be exhaustive, and I don't have any systematic account of such principles. Nevertheless, it seems clear that any plausible, informative principle of this sort is going to need to appeal, as do P1-P4, to some objective normative notion, such as, the notion of what an

<sup>35</sup> If we let  $\varphi$  range over negative acts (i.e., acts of refraining from performing some positive act) as well as positive acts, we would get the counterintuitive result that Sally is subjectively required to refrain from closing Gate 3 given that she knows that she is objectively required to refrain from closing Gate 3.

Also, note that given this principle there is a sense in which objective norms *are* action-guiding. After all, the belief that one is objectively required to  $\varphi$  will, given P2, lead the normatively conscientious agent to believe that she is subjectively required to  $\varphi$ . And this, presumably, will lead her to  $\varphi$ . So even if our first-person practical deliberations are not directly based on our judgments about what we are objectively required to do, it will sometimes be indirectly based on such judgments.

agent has objective reason to do, the notion of what an agent objectively ought to do, or the notion of what an agent is objectively required to do. Those that don't appeal to any such objective normative notion will, I believe, turn out to be either implausible or uninformative.<sup>36</sup> To illustrate the former vice, take the following principle:

**P5** It is subjectively permissible for an agent to  $\varphi$  if and only if her  $\varphi$ -ing would maximize expected utility.

This principle is implausible and not just because (objective) utilitarianism is implausible. Even if we were to grant (objective) utilitarianism—the view according to which it is (objectively) permissible for an agent to  $\varphi$  if and only if her  $\varphi$ -ing would maximize utility, P5 would still be implausible and precisely because it fails to appeal to the agent's (reasonable) beliefs about objective normative notions.

To illustrate, suppose that Peter is considering whether to murder Smith. Assume that Peter's murdering Smith would maximize both actual and expected utility, as Smith's organs could then be used to save Jones. But assume that Peter's refraining from murdering Smith would also maximize both actual and expected utility, for the actual and expected utility in saving Jones would only just offset the actual and expected disutility in murdering Smith. Assume, then, that the two options are tied for first-place with regard to both actual and expected utility. Assume that Peter justifiably believes there is no chance that his refraining from murdering Smith is objectively permissible and that there is some subjective chance (whatever subjective chance there is that utilitarianism is false) that murdering Smith is objectively impermissible. Assume that Peter's refraining from murdering Smith is no more personally costly for Peter than his murdering Smith. Lastly, assume that Peter is in "the normal

<sup>36</sup> I'm concerned with only principles that are universally quantified over agents, for it seems to me that, qua moral philosopher, that's the sort of principles that I should be providing. And all the principles that I discuss here are implicitly universally quantified over agents. Nevertheless, I admit that one could come up with plausible and informative principles about what some particular agent who has certain specific credences regarding the relevant objective normative facts subjectively ought to do and that such a principle needn't explicitly refer to any objective normative notion. But note that one would need to appeal to the agent's credences regarding the relevant objective normative facts when constructing such a principle.

frame of mind" and thus is morally blameworthy (and/or rationally criticizable) for any subjectively impermissible acts that he performs.<sup>37</sup>

In this case, P5 implies that Peter's murdering Smith is subjectively permissible and, thus, not blameworthy. But, surely, given Peter's reasonable uncertainty with respect to utilitarianism's correctness, Peter would be blameworthy for murdering Smith. In murdering Smith, Peter would be taking a big subjective risk of doing what is objectively impermissible. And there is no good reason for Peter to take such a risk. After all, his refraining from murdering Smith carries no subjective risk of doing what is subjectively impermissible. And his refraining from murdering Smith is no more personally costly for him than his murdering Smith. Thus, it seems that P3, not P5, gives the most plausible verdict in this sort of case. Whereas P5 implausibly implies that Peter's murdering Smith is subjectively permissible, P3 rightly implies that Peter's murdering Smith is subjectively impermissible. So, given the plausibility of principles such as P3, it seems that principles, such as P5, that seek to provide substantive answers about what agents subjectively ought to do without appealing to any objective normative notion are going all turn out to be implausible.

As I just suggested, the point about the implausibility of P5 generalizes. To illustrate, suppose that F is some feature of actions that doesn't appeal to any objective normative notion. Suppose, for instance, that 'F' stands for 'an act that would save a person's life'. Any principle that holds that an agent subjectively ought to  $\varphi$  if and only if  $\varphi$  is F (or if and only if the agent believes that  $\varphi$  is F) is going to be implausible precisely because it doesn't appeal to the agent's beliefs about some objective normative notion. Unless the agent believes that she objectively ought to perform acts that are F, it can't be that she subjectively ought to perform acts that are F (or acts that she believes that are F). For if she, instead, believes that she objectively ought to refrain from performing acts that are F, then it will be false that she subjectively ought to perform acts that are F (or acts that she believes that are F). So, again, we find that any plausible, informative principle about what an agent subjectively ought to do must appeal to the agent's beliefs about some objective normative notion.

<sup>37</sup> Again, I'm assuming that, roughly speaking, an act is subjectively wrong if and only if an agent in the normal frame of mind would be blameworthy for performing it—see GIBBARD 1990, pp. 44-45.

Of course, I would admit that the following sorts of principles are plausible even though none of them appeal to any objective normative notion:

- P6 It is subjectively permissible for an agent to  $\varphi$  if and only if it is not subjectively impermissible for her to  $\varphi$ .
- P7 An agent subjectively ought to  $\varphi$  if and only if she has most subjective reason to  $\varphi$ .
- P8 An agent is subjectively required to  $\varphi$  if and only if she has decisive subjective reason to  $\varphi$ .

But although these principles are plausible, they are also completely uninformative. They don't provide any substantive answers regarding what agents subjectively ought to do. Thus, it seems that all principles about what agents subjectively ought to do that fail to appeal to some objective normative notion turn out to be either implausible in the way that P5 is or uninformative in the way that P6-P8 are.

It seems, then, that objective normative notions are more fundamental than their subjective counterparts, for we can arrive at plausible, substantive answers about the latter only by appealing to the former. What's more, it seems that, given P4, we have a prima facie obligation to do what we can to dispel as much of this normative uncertainty regarding these objective normative notions as possible by theorizing both about what we have objective reason to do and about what we objectively ought to do. This is what I aim to do in this book, and it is, as I've just shown, a project of fundamental importance.

### **§1.5 Conventions that I'll follow throughout the book**

Sections, subsections, and tables will be numbered according to the chapter in which they appear. Thus, §3.2 is the second section of Chapter 3, §§3.2.1 is the first subsection of §3.2, and Table 4.1 is the first table in Chapter 4. Major propositions will initially be set off from the rest of the text by additional line spacing above and below. These propositions will either be given an abbreviation, such as 'MR' for 'moral rationalism', or be numbered according to the chapter in which they appear. Thus, 2.11 refers to the eleventh numbered proposition of Chapter 2. These numbers and abbreviations will appear in boldface type in their first appearance. The use

of lowercase letters will indicate specific instances of a more general proposition. Thus, 4.1a and 4.1b are both instances of 4.1.

References in the text and footnotes to works cited in the bibliography are given by the author's last name, appearing in small capitals, and the year of publication. For instance, 'SCANLON 1998' refers to Thomas Scanlon's *What We Owe to Each Other*, which was published in 1998. See the bibliography for complete bibliographic information.

The act-tokens available to a given subject, *S*, will be labeled  $a_1, a_2, \dots, a_n$ . Their corresponding outcomes will be labeled  $o_1, o_2, \dots, o_n$ . Times will be labeled  $t_0, t_1 \dots t_n$ , where  $t_0$  refers to the present. Since acts are rarely, if ever, performed in an instant, we should assume that, when I write, say, 'S performed  $a_1$  at  $t_3$ ', ' $t_3$ ' refers to the interval of time over which S performed  $a_1$ .

An agent's act alternatives consist in the set of mutually exclusive and jointly exhaustive *act-tokens* that are available to her at a given time.<sup>38</sup> This will always include the act of "doing nothing"—that is, the act of intentionally remaining motionless. The act of refraining from performing some act (type or token) is never itself an act-token but is always an *act-type*, which can be instantiated by numerous act-tokens.<sup>39</sup> I'll refer to the "act" of refraining from performing a specific act using the tilde (~). Thus, 'S performs  $\sim a_1$ ' stands for 'S refrains from performing  $a_1$ '. I'll use  $A_1, A_2, \dots, A_n$  to designate act-types, such as lying, cheating, stealing, etc. These consist in sets of all and only those act-tokens that fall under the relevant descriptions. I'll use  $O_1, O_2, \dots, O_n$  to designate the sets of outcomes corresponding to all and only those act-tokens in  $A_1, A_2, \dots, A_n$ , respectively.

In addition to act-tokens and act-types, there are also *act-sets*. An act-set is a set of one or more acts (types or tokens) that are all jointly performable by a single agent. The acts in the set may be simple or compound, synchronous or asynchronous, consecutive or inconsecutive. I'll use  $\alpha_1, \alpha_2, \dots, \alpha_n$  to designate act-sets. I'll use ' $a_i$ ', ' $o_i$ ', ' $t_i$ ', ' $\alpha_i$ ', ' $A_i$ ', and ' $O_i$ ' as variables ranging over act-tokens, outcomes, times, act-sets, act-types, and sets of outcomes, respectively. Italicized letters, such as  $x, y$ , and  $z$ , will be used as variables that range over individual acts (be they types or tokens). I'll sometimes use  $x_1, x_2, \dots, x_n$  to designate distinct acts. Greek

<sup>38</sup> An act-token is a particular act performed by a particular agent at a particular time. Act-tokens are performable at most once and are not the sort of thing that can be instantiated.

<sup>39</sup> Act-types are universals that can, in general, be instantiated by a number of distinct act-tokens.

letters, such as  $\varphi$ ,  $\chi$  and  $\psi$ , range over everything for which agents can have reasons, including beliefs, desires, act-tokens, act-types, and act-sets.

Technical terms and phrases will be defined as they come up and will be set off by italics in their initial appearance. Readers should consult the glossary if they ever need to be reminded of a definition.

I will typically use 'she', 'her', and 'hers' as generic personal pronouns whenever the gender of the referent is unspecified, although I will occasionally use 'he', 'him', and 'his' as generic personal pronouns, as when I must do so in order to remain consistent with some quoted passage. I realize that this is far from ideal, but all the alternatives seem to me to be just as bad if not worse.

## CHAPTER 2

---

# Consequentialism and Moral Rationalism<sup>1</sup>

- 2.1 The too-demanding objection: How moral rationalism leads us to reject utilitarianism
- 2.2 The argument against utilitarianism from moral rationalism
- 2.3 How moral rationalism compels us to accept consequentialism
- 2.4 What is consequentialism?
- 2.5 The presumptive case for moral rationalism
- 2.6 Some concluding remarks

IN THIS CHAPTER, I make a presumptive case for *moral rationalism*: the view that agents can be morally required to do only what they have decisive reason to do, all things considered.<sup>2</sup> I argue that this view compels us to accept act-consequentialism, but at the same time leads us to reject all traditional versions of the theory. I begin by explaining how moral rationalism leads us to reject utilitarianism.

### §2.1 The too-demanding objection: How moral rationalism leads us to reject utilitarianism

As we saw in Chapter 1, utilitarianism is too demanding. It implies that agents are morally required to sacrifice their projects, interests, and special relationships whenever doing so would produce more, even just slightly more, aggregate utility than not doing so would. Thus, according to utilitarianism, I'm morally required to sacrifice my life, to neglect my relationship with my daughter, and to abandon my project of completing this book if I could thereby produce more, even just slightly more, aggregate utility. To demand that I make such sacrifices for the sake of such miniscule gains in aggregate utility is to demand more from me than can be rightfully or reasonably demanded of me.

<sup>1</sup> Much of the material in this chapter is drawn from PORTMORE 2011.

<sup>2</sup> I assume, contrary to COPP 1997, that there is a normative standpoint from which we can judge what an agent has decisive reason to do, all things considered—in other words, that there is a normative standpoint from which we can judge what an agent just plain ought to do. See MCLEOD 2001 for a reply to Copp.

To say that a given theory is *too* demanding is not merely, or even necessarily, to say that it demands quite a lot from agents in certain circumstances. After all, almost all moral theories demand quite a lot from agents in at least some circumstances.<sup>3</sup> What's more, a theory can be too demanding in part because some of its demands, though quite small, are more than can be rightfully or reasonably demanded of agents. Such is the case with utilitarianism.

To illustrate, imagine that we live in a utopian world in which everyone is not only materially well-off but also extremely happy. And assume, as is in fact the case, that in addition to two other computers I own an ultra-portable laptop computer that I use only for travel. This computer is admittedly a luxury, which I use to pass the time on planes and in airports. To demand that I forfeit this luxury for the sake of promoting the greater good is not to impose any great demand on me. Yet a theory that required me to hand over my computer to some even more well-to-do person who already has *four* computers, including one ultra-portable laptop, would be too demanding. It's one thing to think that, in a non-utopian world, I would be required to give up such a luxury for the sake of saving some poor child's life, but it's quite another to think that, in a utopian world, I would be required to give up such a luxury for the sake of providing some well-to-do person with even more luxury than I have. Nevertheless, utilitarianism requires me to do just that if that's what would maximize utility, as where, for instance, this well-to-do person would get slightly more enjoyment out of my laptop than I would. And this is true even if the net gain to be had is but one measly utile, which is, as you'll recall, equivalent to someone's experiencing the mildest of pleasures for the briefest of moments. To demand that I make such a sacrifice for the sake of benefiting someone who is even better off than me is to demand more from me than can rightfully or reasonably be demanded of me.

The foregoing suggests that there are two possible senses in which utilitarianism might be too demanding: (1) it holds that agents are morally required to make sacrifices that they are not, in fact, morally required to make or (2) it holds that agents are morally required to make sacrifices that they do not have decisive reason to make, all things considered.<sup>4</sup> For now, let's call these sense<sub>1</sub> and sense<sub>2</sub>, respectively. The claim that utilitarianism is too demanding in sense<sub>1</sub> is, perhaps, the more common of the two, but it

<sup>3</sup> Paul Hurley (2006, p. 681) makes this point as well.

<sup>4</sup> In thinking about utilitarianism's being too demanding in sense<sub>2</sub>, I have been greatly influenced by DORSEY 2009.



is certainly not the only sense in which utilitarianism is claimed to be too demanding. To see this, consider that some utilitarians concede that their theory is too demanding but deny that this constitutes an objection to it.<sup>5</sup> If their claim was that utilitarianism is too demanding in sense<sub>1</sub>, then this would make no sense, for if utilitarianism holds that we are morally required to make sacrifices that we are not, in fact, morally required to make, then it would indeed be false. One can, however, admit that utilitarianism is too demanding in sense<sub>2</sub> and yet deny that the theory is false. The thought would be that morality is itself too demanding, and so even if utilitarianism is too demanding, it could, nonetheless, be the correct moral theory.

Samuel Scheffler makes the point quite eloquently:

Suppose that morality as represented by a certain philosophical theory strikes us as too demanding in what it requires of people. Does this, in principle, provide the basis for a legitimate objection, either to the theory or to morality itself? If so, is the problem that the theory has distorted the content of morality? Or is it that morality itself is excessively demanding, so that while the theory may be an accurate representation of the content of morality, people have reason to treat moral considerations as less weighty or authoritative than we may previously have supposed? (1992, 5)

But what does it mean to say that morality is itself too demanding? Here, the claim would seem to be that morality is too demanding in sense<sub>2</sub>. As David Sobel puts it, “What morality asks, we could say, is too much to be the thing the agent has most reason to do all things considered, but not too much to count as what morality asks” (2007, p. 14). So, clearly, there are these two senses in which a moral theory might be thought to be too demanding. Let us call those theories that are too demanding in sense<sub>1</sub> *erroneously demanding* and call those theories that are too demanding in sense<sub>2</sub> *unreasonably demanding*.

I want to focus exclusively on the idea that utilitarianism is unreasonably demanding, for this is something that most philosophers

<sup>5</sup> See, for instance, SIDGWICK 1966, SINGER 1999 (pp. 289, & 308-9), and D. SOBEL 2007 (p. 14). To be fair, I should note that Sobel never officially endorses consequentialism, let alone utilitarianism specifically. Nevertheless, we can glean from what he says that he thinks both that utilitarianism is too demanding in sense<sub>2</sub> and that this is no objection to the view. Interestingly, Paul Hurley has argued that Peter Singer is, despite his explicit denials of MR, implicitly committed to MR (or at least something very close to it)—see HURLEY 2009.

(utilitarians and non-utilitarians alike) can agree on.<sup>6</sup> Given any plausible conception of practical reasons, utilitarianism is unreasonably demanding, for agents often do not have decisive reason to make the sorts of sacrifices that utilitarianism requires them to make. For instance, even if killing myself would maximize aggregate utility, it's unlikely that I would have decisive reason to do so if the net gain in aggregate utility would be only one util. Of course, as I've noted above, one can admit that utilitarianism is unreasonably demanding and deny that this constitutes an objection to the theory. But some do take the fact that utilitarianism is unreasonably demanding as an objection to the theory.<sup>7</sup>

What, then, accounts for the disagreement? It seems to be this: unlike those who deny that a theory's being unreasonably demanding constitutes an objection to that theory, those who think that it does accept the following thesis, which I call *moral rationalism*:<sup>8</sup>

**MR** If a subject, *S*, is morally required to perform an act, *x*, then *S* has decisive reason to perform *x*, all things considered.<sup>9</sup>

<sup>6</sup> Utilitarians that would agree with this include Henry Sidgwick (1966) and Peter Singer (1999, pp. 289, & 308-9). Non-utilitarians that would agree with this include Paul Hurley (2006) and Sarah Stroud (1998).

<sup>7</sup> For instance, Paul Hurley has, in a recent paper, argued against traditional forms of consequentialism, and these arguments proceed "within the widely held framing conviction that any plausible theory of moral standards must be such that rational agents typically have decisive reasons to avoid what these standards identify as cases of wrongdoing" (2006, p. 704). Sarah Stroud argues against consequentialism on the grounds that it conflicts with the following thesis: "If *S* is morally required to  $\phi$ , then *S* has most reason to  $\phi$ "—see STROUD 1998 (esp. p. 171 and pp. 182-184).

<sup>8</sup> The thesis that I call *moral rationalism* sometimes goes by other names. David Brink calls MR "the supremacy thesis" (1997, p. 255), Stephen Darwall calls it "supremacy" (2006b, p. 286), Samuel Scheffler calls it "the claim of overridingness" (1992, pp. 52-54), John Skorupski calls it "the principle of moral categoricity" (1999, p. 170), Sarah Stroud calls it the "overridingness thesis" (1998, p. 171), and R. Jay Wallace calls it the "optimality thesis" (2006, p. 130).

<sup>9</sup> The essential idea is that agents can be morally required to do only what they are rationally required to do. Now, some philosophers (see, e.g., STROUD 1998) formulate this thesis using 'most reason' in place of 'decisive reason'. This formulation is equivalent to mine unless either some reasons have no (rational) requiring strength (see GERT 2004 and DANCY 2004) or some reasons with requiring strength are silenced, undermined, or bracketed off by other factors or considerations (see, e.g., SCANLON 1998). If either is the case, then *S*'s having most reason to perform *x* would not entail *S*'s being rationally required to perform *x*. And, in that case, I assume that Stroud and the others would want to replace 'most reason' with 'decisive reason', where '*S* has decisive reason to perform *x*' just means '*S*'s reasons are such as to make *S* (objectively) rationally required to perform *x*'. In

If, on the one hand, MR is true, then the correct moral theory cannot be unreasonably demanding. And this explains why those who accept MR reject utilitarianism on account of its being unreasonably demanding, for MR and the fact that utilitarianism is unreasonably demanding implies that the theory is erroneously demanding as well. But if, on the other hand, MR is false, then the correct moral theory can be unreasonably demanding without being erroneously demanding. And this explains why those who reject MR are content with the fact that utilitarianism is unreasonably demanding.

So far, we've seen how the observation that utilitarianism is unreasonably demanding will lead those who accept MR to reject utilitarianism. Of course, I haven't yet provided any reason to think that MR is true. In §2.5, I'll rectify this by making a presumptive case for MR. But before proceeding with that, I want to show how the objection that utilitarianism is unreasonably demanding is really just a species of a much more general objection.

## **§2.2 The argument against utilitarianism from moral rationalism**

The idea that utilitarianism is objectionable insofar as it is unreasonably demanding is really just a species of a much more general objection, for the objection proceeds on the assumption that MR is true, and once we accept MR, we should object to utilitarianism whenever it requires us to perform acts that we don't have decisive reason to perform, whether these acts be self-sacrificing acts or even self-benefiting acts. Thus, we should object to utilitarianism not only on the grounds that it sometimes requires agents to make sacrifices that they don't have decisive reason to make, but also on the grounds that it sometimes requires agents to perform self-benefiting acts that they don't have decisive reason to perform.

To illustrate, let's suppose that a woman named Anita would benefit from murdering her estranged uncle. Assume that she's crafty enough to get away with it and that, if she did, she would finally receive the heart transplant that she so desperately needs, for her uncle is an organ donor with the same rare blood group and tissue type that she has. Let's further suppose that murdering her uncle would maximize aggregate utility, for let's assume that the gain in utility resulting from Anita's heart transplant would slightly more than offset the loss in utility resulting from her uncle's

other words, S has decisive reason to perform  $x$  if and only if S lacks sufficient reason to perform any alternative to  $x$ .

murder. Given this further stipulation, utilitarianism requires Anita to benefit herself by murdering her uncle. Arguably, though, she does not, in this instance, have decisive reason to commit murder. It seems that, in this instance, the moral reason that she has to refrain from committing murder outweighs the self-interested reason she has to do so.

So, if we accept MR, we should object to utilitarianism because it sometimes requires agents to perform self-benefiting acts that they don't have decisive reason to perform. We can, then, offer the following general schema for constructing an argument against utilitarianism on the basis of MR:

- 2.1 There exists an act,  $x_1$ , available to S that would maximize aggregate utility but that S does not have decisive reason to perform, all things considered.
- 2.2 If utilitarianism is true, then S is morally required to perform any act that would maximize aggregate utility. (From the definition of 'utilitarianism')
- 2.3 So, if utilitarianism is true, then S is morally required to perform  $x_1$ . (From 2.1 and 2.2)
- 2.4 S is morally required to perform  $x_1$  only if S has decisive reason to perform  $x_1$ , all things considered. (From MR)
- 2.5 So, S is not morally required to perform  $x_1$ . (From 2.1 and 2.4)
- 2.6 Therefore, utilitarianism is not true. (From 2.3 and 2.5)

The argument is deductively valid, and the only controversial premises are 2.1 and 2.4. Since 2.4 follows from MR and since I'm interested in what follows if MR is true, I'll focus on 2.1. To see just how plausible 2.1 is, consider the following two plausible candidates for  $x_1$ .

First, let  $x_1$  be the act of your sacrificing the life of your partner (whose life, we'll assume, is well worth living) for the sake of saving some stranger's life, where your doing so would maximize aggregate utility, but where the net gain in aggregate utility would be quite small: one util. Suppose, for instance, that both lives contain nearly the same amount of utility but that what accounts for there being more utility in saving the stranger's life is the fact that her life contains slightly more utility: an extra second of mild pleasure. Assume that, but for this difference, the utility of your sacrificing the life of your partner would be the same as that of your refraining from doing so.

Clearly, in this case, you lack decisive reason to save the stranger's life, for the reason that you have to save your partner is at least as strong as, if

not stronger than, the reason you have to bring about an extra second of mild pleasure. After all, let's assume that you have a very special relationship with your partner, involving years of personal history together and mutual love and respect for one another. Let's even assume that you care more about the welfare of your partner than you do about the welfare of this stranger, that you care more about the welfare of your partner than you do about producing small net gains in aggregate utility, and that saving your partner is what would best serve your own interests. On what plausible theory of practical reasons, then, would there not be at least sufficient reason for you to save your partner instead of this stranger? I can't think of any.<sup>10</sup> Thus, this seems to be a clear case where a subject, S (viz., you), could maximize aggregate utility by performing an act,  $x_1$  (specifically, the act of saving the stranger instead of your partner), but where  $x_1$  is not something S has decisive reason to do.

Second, let  $x_1$  be the act of murdering your partner so as to ensure that each of a billion and one people get some minor benefit. Assume that, if you murder your partner, the loss in utility would be one billion utiles, and assume that the benefit to each of the billion and one people would be one utile. Thus, the net gain to be had in your murdering your partner is but one utile. To make 2.1 even more plausible, assume that your partner is one of society's worst off and undeservedly so and that the billion and one people who you could instead benefit are society's most well-to-do and undeservedly so. In fact, let's assume that they are well-to-do only as a result of exploiting society's worst off, including your partner.

The idea that, in this situation, all the various reasons that you have to refrain from murdering your partner are, even when taken together, insufficient to successfully counter the fairly weak reason that you have to produce a net gain of one utile is quite implausible. Consider all the reasons that you have to refrain from murdering your partner: (1) the reason that you have to respect your partner's autonomy, (2) the reason that you have to give priority both to your own welfare and to the welfare of those whom you most love, (3) the reason that you have to give priority to society's worst off, and (4) the reason that you have to give priority to

<sup>10</sup> I can, of course, think of an *implausible* theory of practical reason according to which you would not have even sufficient reason to save your partner: that is, the theory according to which S has sufficient reason to perform  $x$  if and only if S's performing  $x$  would maximize aggregate utility. But I know of no one who has argued for such a theory of practical reason. Even utilitarians who think that there is always sufficient reason to maximize aggregate utility admit that there is also often sufficient reason to perform acts that don't maximize aggregate utility—see, for instance, SIDGWICK 1966.

the deserving over the undeserving. Given how much reason you have not to murder your partner and how little reason you have to produce a small net gain in aggregate utility, we should deny that you have decisive reason to murder your partner.

Given these two plausible candidates for  $x_1$ , we can be confident that there are at least two acts that would have both the property of being what would maximize aggregate utility as well as the property of not being what the agent has decisive reason to do, all things considered. So we should accept 2.1, and, therefore, acknowledge that utilitarianism cannot be the correct moral theory if MR is true. The problem for utilitarianism vis-à-vis MR is that utilitarianism ignores two important classes of rather weighty reasons—reasons that, in certain instances, have sufficient weight to decisively oppose the reason that one has to maximize utility: (1) reasons that have nothing to do with *promoting* utility, such as the reason one has to ensure that utility is fairly distributed and (2) reasons that stem from the special relations that we bear to ourselves and our loved ones, such as the reason that we have to favor ourselves and our loved ones when deciding whose utility to promote. If a theory of morality, such as utilitarianism, ignores such reasons, it will conflict with MR.<sup>11</sup>

And it's not just utilitarianism that ignores whole classes of reasons. All forms of traditional act-consequentialism do so as well. According to all forms of *traditional act-consequentialism* (TAC), we should accept the following *criterion of rightness* (cr):

**TAC<sub>cr</sub>** An act is morally permissible if and only if it maximizes the good (impersonally construed).

TAC<sub>cr</sub> ignores two important classes of rather weighty reasons—reasons that, in certain instances, have sufficient weight to decisively oppose the reason that one has to maximize the good: (1) reasons that have nothing to do with promoting the good, such as the reason one has to refrain from violating someone's autonomy even when doing so is a means to promoting the good, and (2) reasons that stem from the special relations that we bear to ourselves and our loved ones, such as the reason one has to promote the good by saving one's own loved one as opposed to by helping some stranger save her loved one.<sup>12</sup>

<sup>11</sup> This sort of argument comes from STROUD 1998.

<sup>12</sup> Even if the traditional act-consequentialist holds, say, that there is more value in a person's life being saved by a loved one than there is in a person's life being saved by some

Insofar as we think either that we have reasons to do things besides promote the good or that the weight of our reasons to perform various good-promoting acts depends on our relationship to those whose good would thereby be promoted, TAC<sub>cr</sub> fails to acknowledge certain reasons and/or their proper weights.<sup>13</sup> And this means that TAC<sub>cr</sub> will conflict with MR. Indeed, we could easily revise the argument that I gave against utilitarianism above so as to constitute an argument against traditional act-consequentialism. We need only substitute 'traditional act-consequentialism' for 'utilitarianism' and 'the good' for 'aggregate utility' in the above argument. I leave this as an exercise for the reader.

### §2.3 How moral rationalism compels us to accept consequentialism

So we've seen how moral rationalism leads us to reject traditional versions of act-consequentialism, such as utilitarianism.<sup>14</sup> Interestingly, though, moral rationalism also compels us to accept act-consequentialism (hereafter simply 'consequentialism'). Or so I will argue. I'll argue that at the heart of utilitarianism lies a certain conception of practical reasons that when conjoined with MR compels us to accept consequentialism.

Many utilitarians and non-utilitarians agree that there is something compelling, perhaps, even "spellbinding," about utilitarianism and that this is what explains: (1) why utilitarianism has persevered despite its implications being so wildly at odds with our most firmly held moral convictions; (2) why it tends "to haunt even those of us who will not believe in it" (FOOT 1985, p. 196); and (3) why "the move to rule utilitarianism seems to be an unsatisfactory answer to the problem of reconciling utilitarianism with common moral opinion" (FOOT 1985, p. 198).

But what is utilitarianism's compelling idea? As many philosophers see it, it is consequentialism itself (e.g., FOOT 1985, p. 196). And, as I see it, what compels us to accept consequentialism is the fact that it embodies a

stranger, she must deny that there is, other things being equal, any gain in value to be had by saving one's own loved as opposed to helping someone else to save her loved one.

<sup>13</sup> I borrow this point from STROUD 1998. Paul Hurley (2009) makes a similar point, although he focuses on the fact that TAC<sub>cr</sub> ignores "non-impersonal reasons." I think, though, that the challenge is broader than that, for it seems that a person could consistently hold that there are impersonal reasons that have nothing to do with promoting the good.

<sup>14</sup> The reader should assume that, unless otherwise indicated, 'consequentialism' is short for 'act-consequentialism' and that 'nonconsequentialism' is short for 'non-act-consequentialism'. On my usage, then, rule-consequentialism is a version of nonconsequentialism.

certain conception of practical reasons: what's sometimes called the '*teleological* conception of *reasons*' (SCANLON 1998, p. 84). On this view, an agent's reasons for performing an action are determined by her reasons for preferring its outcome to those of the available alternatives, such that, if S has most reason to perform *x*, then, of all the outcomes that S could bring about, S has most reason to desire that *x*'s outcome obtains.<sup>15</sup> The intuitive idea is that since our actions are the means by which we affect the way the world goes, and since our intentional actions necessarily aim at making the world go a certain way, it is only natural to suppose that our reasons for action are a function of our reasons for preferring some of these possible worlds to others, such that what each agent has most reason to do is to bring about the possible world, which of all those available to her, is the one that *she* has most reason to want to obtain.

Of course, we must acknowledge that no conception of reasons could ever compel us to accept consequentialism unless MR is true. If MR is false, then moral requirements do not entail requirements of reason. And if moral requirements do not entail requirements of reason, then it's hard to see why we should think that any conception of reasons could compel us to accept consequentialism. It seems, then, that it is actually not a single idea, but the following argument, that compels us to accept consequentialism:<sup>16</sup>

- 2.7 An act's deontic status is determined by the agent's reasons for and against performing it, such that, if a subject, S, is morally required to perform an act, *x*, then S has most reason to perform *x*.<sup>17</sup> (Moral Rationalism)

<sup>15</sup> Although this view is one that many find intuitively compelling, it is not without its critics—see, for instance, SCANLON 1998 (pp. 50-64 & chap. 2). But see Chapter 3 and PORTMORE 2009a for a defense of this view.

<sup>16</sup> In earlier works, I had misidentified consequentialism's compelling idea first as the "idea that it is always permissible to bring about the best available state of affairs" (PORTMORE 2005, p. 95) and then subsequently as the idea that it is always permissible to act so as to bring about the outcome that one has reason to prefer above all other available alternatives (PORTMORE 2007, p. 50). Both are problematic—see SCHROEDER 2007 and SACHS 2009, respectively.

<sup>17</sup> As mentioned in note 9 above, the second clause in 2.7 is equivalent to MR assuming both that all reasons have (rational) requiring strength and that no reason with requiring strength is ever silenced, undermined, or bracketed off by other factors or considerations. If either assumption is false, I'll need to revise this argument so as to replace 'most reason' with 'decisive reason' throughout, which is something I could do with no ill effect for my argument. I choose, here, to talk about what there is most reason to want and to do only because this talk seems more natural given that I'm making comparative judgments.



- 2.8 The agent's reasons for and against performing an act are determined by her reasons for and against preferring its outcome to those of the available alternatives, such that, if S has most reason to perform  $x$ , then, of all the outcomes that S could bring about, S has most reason to desire that  $x$ 's outcome obtains.<sup>18</sup> (The Teleological Conception of Practical Reasons)
- 2.9 Therefore, an act's deontic status is determined by the agent's reasons for and against preferring its outcome to those of the available alternatives, such that, if S is morally required to perform  $x$ , then, of all the outcomes that S could bring about, S has most reason to desire that  $x$ 's outcome obtains.<sup>19</sup> (Consequentialism)

## §2.4 What is consequentialism?

I've been assuming that 'outcome' and 'consequentialism' are to be construed broadly. Broadly construed, an act's outcome is the possible world that would be actual if it were performed.<sup>20</sup> And, broadly construed, consequentialism is the view that an act's deontic status is determined by how its outcome ranks relative to those of the available alternatives on some evaluative ranking, such that an act is morally obligatory only if its outcome outranks those of the available alternatives.<sup>21</sup> But even so, 2.9 is

The first clause in 2.7 goes beyond MR itself in positing an explanatory relation between an act's deontic status and the reasons there are to perform it. I'm assuming, then, that moral rationalists would be happy to endorse this.

<sup>18</sup> Michael Smith (2006) has argued for the teleological conception of reasons, although he formulates it a bit differently: "(x)(t)(x at t has all things considered reason to  $\varphi$  in circumstances C iff  $\varphi$ -ing is the unique action of those x can perform at t that brings about what x would desire most happens in C if he had a desire set that satisfied all requirements of...reason."

<sup>19</sup> It may be that some outcomes are incommensurable such that, in some instances, there is an alternative to  $x$ 's outcome that S has neither more nor less, nor equal, reason to desire. In such instances, S will not have most reason to desire that  $x$ 's outcome obtains and, consequently, will not be morally required to perform  $x$ .

<sup>20</sup> On why consequentialists should construe an act's outcome to include not merely its causal consequences but everything that would be the case were it to be performed, see BROOME 1991 (pp. 3-4), SCHEFFLER 1994 (pp. 1-2), SOSA 1993 (pp. 101-2), WILLIAMS 1973 (pp. 86-87), and especially CARLSON 1995 (p. 10 & chap. 4).

<sup>21</sup> I write 'only if' as opposed to 'if and only if' so as not to exclude satisficing versions of consequentialism, which deny that it is always obligatory to produce the highest ranked outcome. I define 'consequentialism' in terms of 'some evaluative ranking' as opposed to a ranking of outcomes with respect to their goodness so as not to exclude egoism, which ranks outcomes not with respect to how good they are, but with respect to how good they

not obviously equivalent to this definition. According to this definition, consequentialist theories hold that acts are permissible or impermissible in virtue of how their outcomes rank relative to those of the available alternatives on some evaluative ranking. According to 2.9, by contrast, consequentialist theories hold that acts are permissible or impermissible in virtue of the agent's reasons for and against preferring its outcome to those of the available alternatives. What, then, is the connection between an evaluative ranking of outcomes and a ranking of outcomes in terms how much reason the agent has to want each of them to obtain? The answer lies in understanding what an evaluative ranking of outcomes is. It is a ranking of outcomes in terms of the agent's reasons (or some subset of her reasons) for preferring each outcome to the available alternatives, for it seems that what ties all the different varieties of goodness—such as *goodness for* or *goodness that*—together is that they are all to be understood in terms of reasons (or some subset reasons) for desiring.

For instance, we can understand *better that* and *better for* as follows:<sup>22</sup>

***Better that:*** For all states of affairs *p* and *q*, it is better that *p* is the case than that *q* is the case if and only if the set of all the right kind of reasons to prefer its being the case that *p* to its being the case that *q* is weightier than the set of all the right kind of reasons to prefer its being the case that *q* to its being the case that *p*.<sup>23</sup>

***Better for:*** For all subjects *S* and all states of affairs *p* and *q*, it is better for *S* that *p* is the case than that *q* is the case if and only if the set of all the right kind of reasons to prefer, for *S*'s sake, its being the case that *p* to its being the case that *q* is weightier than the set of all the right kind

are for the agent. And, strictly speaking, consequentialism needn't make the deontic statuses of actions a function of some single evaluative ranking; it could instead make it a function of some set of evaluative rankings, as is the case with dual-ranking act-consequentialism—see Chapter 1 and especially Chapter 5.

<sup>22</sup> These are adapted from SCHROEDER 2008a and SCHROEDER 2008b.

<sup>23</sup> If facts such as the fact that an evil demon will torture me unless I prefer its being the case that *p* to its being the case that *q* are genuine reasons for me to have this preference and not just reasons for me to desire, and to act so as to ensure, that I have this preference, then the right kind of reasons must exclude such reasons, for even if such a fact did constitute a genuine reason for me to prefer its being the case that *p* to its being the case that *q*, it would have no bearing on whether or not it is better that *p* is the case than that *q* is the case. See PARFIT 2008, RABINOWICZ AND RØNNOW-RASMUSSEN 2004, SCHROEDER 2008a, and ZIMMERMAN 2009 for more on this problem, which is known as *the wrong-kind-of-reasons problem*.

of reasons to prefer, for *S*'s sake, its being the case that *q* to its being the case that *p*.<sup>24</sup>

Different consequentialist theories employ different evaluative rankings, which, in some instances, appeal to different varieties of goodness. For instance, on egoism, outcomes are ranked in terms of how *good* they are *for* the agent, whereas, on traditional act-consequentialism, outcomes are ranked in terms of how *good* it would be *that* they obtain.<sup>25</sup> And since all these varieties of goodness are ultimately to be understood in terms of reasons for preferring, we should understand consequentialism to be the view according to which an act's deontic status is determined by the agent's reasons for and against preferring its outcome to those of its alternatives. We should, therefore, take 2.9 to be equivalent to the view that an act's deontic status is determined by how its outcome ranks relative to those of the available alternatives on some evaluative ranking, such that an act is morally obligatory only if its outcome outranks those of all other available alternatives.

If this is right, then what differentiates one consequentialist view from another is their differing commitments concerning what agents have most reason to desire, all things considered. Whereas, say, the utilitarian is committed to the view that each agent has most reason, all things considered, to desire the outcome that she has most impartial reason to desire, the egoist is committed to the view that each agent has most reason, all things considered, to desire the outcome that she has most self-interested reason to desire. Thus, the utilitarian is committed to the view that there are no partial reasons for preferring one outcome to another, or, if there are, that these are always overridden by whatever impartial reasons there are for preferring the one to the other. And the egoist, by contrast, is committed to the view that there are no altruistic reasons for preferring one outcome to another, or, if there are, that these are always overridden by whatever self-interested reasons the agent has for preferring the one to the other.

<sup>24</sup> To avoid circularity, we'll need to find some way of explicating "for *S*'s sake" in terms other than "for *S*'s own good." Chris Heathwood (2008) presents an interesting argument against this sort of account, but see JOHANSSON 2009 for a possible reply. See also RØNNOW-RASMUSSEN 2007 and RØNNOW-RASMUSSEN 2009.

<sup>25</sup> As I understand it, egoism is the view that holds both that act-consequentialism is true and that an act produces the outcome that *S* has the most reason to desire if and only if it maximizes *S*'s utility.

The idea that egoists and utilitarians are committed to certain views about what agents have most reason, all things considered, to desire may seem a bit strange. In particular, it may seem strange to suggest that someone who holds that an act is morally permissible if and only if it maximizes the good could fail to be a consequentialist just because she thinks that partial reasons can provide an agent with decisive reason for preferring a worse outcome to a better outcome. She might, for instance, hold that although it would be better if some child more gifted than her own survives, she ought, all things considered, to prefer that her less gifted child survives. In this case, she will hold that she ought to act so as to ensure that the more gifted child survives even though she ought to prefer the possible world in which she acts instead to ensure that her less gifted child survives. Since she thinks that she is morally required to act so as to produce an outcome that she does not have most reason to desire, she will, if 2.9 is definitive of consequentialism, count as a nonconsequentialist.<sup>26</sup>

So if we take 2.9 to be definitive of consequentialism, then this would seem to imply that many philosophers misuse the term 'consequentialism' to refer to all and only those theories that direct each agent to bring about the best outcome rather than to all and only those theories that direct each agent to bring about the outcome that, of all those available to her, is the one that she has most reason to desire. This is a fairly radical conclusion. I suspect, then, that many readers will reach the opposite conclusion: that I am the one who is misusing the term 'consequentialism'. In fact, none of us are misusing the term, for 'consequentialism' is a term of art. As Walter Sinnott-Armstrong (2006) points out, "In actual usage, the term 'consequentialism' seems to be used as a family resemblance term to refer to any descendant of classic utilitarianism that remains close enough to its ancestor in the important respects. Of course, different philosophers see different respects as the important ones. Hence, there is no agreement on which theories count as consequentialist." My view is that what's most important and compelling about utilitarianism is that it can be seen as deriving from 2.7 and 2.8 given the further assumption that what agents always have most reason to desire, all things considered, is the outcome with the most aggregate utility. Thus, I take 2.9, which is what follows from 2.7 and 2.8 alone, to be definitive of the broader notion: namely,

<sup>26</sup> To say that 2.9 is definitive of act-consequentialism is not to say that 2.7 (i.e., MR) is definitive of consequentialism. It's not. One can accept 2.9 and reject MR, as many traditional act-consequentialists do. This is why MR can lead us to reject traditional versions of act-consequentialism while at the same time compelling us to accept some version of act-consequentialism.

consequentialism. But since others find other aspects of utilitarianism compelling, it's no surprise that they will reject my definition.

We should not quibble, though, over such unimportant terminological issues. If readers find my definition of 'consequentialism' too revisionary, then they should label 2.9 something besides 'consequentialism': 'rational-desire teleology', perhaps.<sup>27</sup> What's important is not what we call 2.9-type theories, but whether such theories are interesting and plausible, as I hope to establish. Of course, I have to call these theories something, and given my views about what's most compelling about a theory like utilitarianism, it makes most sense for me to call them consequentialist theories. Let me just stipulate, then, that, as I'll use the term 'consequentialism', a theory is (act-)consequentialist if and only if it entails 2.9. We can, then, leave aside the rather uninteresting question of whether this is in fact the definition that best comports with common philosophical usage and focus instead on my arguments that consequentialism, so defined, is both interesting and plausible.<sup>28</sup>

## §2.5 The presumptive case for moral rationalism

So far, we've seen how moral rationalism (MR) compels us to accept consequentialism while at the same time leading us to reject all traditional versions of the theory. Of course, these observations are of little interest unless MR is true, and I have yet to provide any reason for thinking that it is. Below, I attempt to rectify this. Let me forewarn the reader, though, that the case that I'll be making in this chapter is, at best, a presumptive one.<sup>29</sup>

<sup>27</sup> Rational-desire teleology would contrast, then, with both *impersonal-value teleology* (the view that an act's deontic status is determined by the impersonal value of its outcome, such that, if S is morally required to perform  $x$ , then S's performing  $x$  would maximize the good) and *personal-value teleology* or egoism (the view that an act's deontic status is determined by the personal value of its outcome, such that, if S is morally required to perform  $x$ , then S's performing  $x$  would maximize S's good).

<sup>28</sup> I find the issue of how 'consequentialism' is to be defined so as to accurately reflect common philosophical usage rather uninteresting. What's more, I think that it may even be pointless to try to resolve the issue, as there doesn't seem to be any set of individually necessary and jointly sufficient conditions that captures all the various uses of the term that are to be found in the literature.

<sup>29</sup> One reason that the case that I'll be making is, at best, a presumptive one is that I don't, for reasons of space, attempt to address all the various objections and counterarguments that have been leveled against MR. For instance, Shaun Nichols (2002) has argued that a recent empirical study, which he himself conducted and which probes people's intuitions about psychopaths, undermines MR. Although I don't have space here to address Nichols's interesting argument, I think that others have adequately defended MR against his

My arguments will establish only that those who reject MR must also reject an intuitively plausible premise. Thus, there is a price to be paid in rejecting MR. Of course, there is often a price to be paid in rejecting a philosophical thesis. And, perhaps, some will think that, in this instance, the price is worth paying.

But even if the arguments in this chapter lack definitiveness, my case for MR does not rest with them alone. Part of my case for MR is that it coheres well with other theses that I'll be arguing for in the book. Ultimately, then, the case for MR depends on how plausible this picture of morality, rationality, and the relationship between the two is on the whole. It won't, however, be until the very last chapter that the whole picture will come fully into view. So, in the meantime, I'll settle for presenting a presumptive argument in favor of MR. But, first, I'll need to explain how MR differs from three related theses about the overridingness of morality and moral reasons. There's potential for confusion here since some philosophers equate MR with the view that morality is overriding, and some even refer to MR as the "overridingness thesis" (e.g., STROUD 1998, p. 171).<sup>30</sup>

Recall that, according to MR, an agent can be morally required to perform an act only if she has decisive reason to perform that act, all things (and, thus, all reasons) considered. It's important not to confuse this thesis about moral requirements and *reasons generally* with the following thesis about moral requirements and *moral reasons specifically*:

**MO** If S has more moral reason to perform  $x$  than to perform  $y$  ( $y \neq x$ ), then S is not morally permitted to perform  $y$ . And, thus, if S

argument. See, for instance, KENNETT 2006. For more on this issue, see also JOYCE 2008a, NICHOLS 2008, and JOYCE 2008b.

<sup>30</sup> The question of whether or not moral requirements override, say, prudential requirements (i.e., requirements of self-interest) is not the same question as whether or not MR is true. Moral requirements override prudential requirements if and only if both (1) there are cases in which moral requirements conflict with prudential requirements and (2), in all those cases, agents have decisive reason, all things considered, to act as they are morally required to act. If ethical egoism, consequentialist prudence, and rational egoism (the view that S has decisive reason to perform  $x$  if and only if S's performing  $x$  would maximize S's utility) were all true, then MR would be true even though moral requirements would not override prudential requirements. In this case, moral requirements would not override prudential requirements for the simple fact that they never conflict.

has most moral reason to perform  $x$ , then  $S$  is morally required to perform  $x$ .<sup>31</sup>

According to MO, moral reasons are *morally overriding* (hence, the abbreviation 'MO'). To say that one type of reason, say,  $m$ -reasons, overrides another, say,  $n$ -reasons, with respect to a certain kind of normative status,  $N$ , is to say that, in any situation in which both types of reasons are present and an act,  $x$ , has a certain  $N$ -status, no modification of that situation that involves affecting only what  $n$ -reasons there are will change  $x$ 's  $N$ -status. That is, if  $m$ -reasons override  $n$ -reasons with respect to an act's  $N$ -status, then even the weakest  $m$ -reason trumps the strongest  $n$ -reason in the determination of that act's  $N$ -status.

To illustrate, suppose that I am morally required to keep my promise to meet with a student to discuss his exam grade, that the reason that I have to keep this promise is a moral reason, that the reason I have to break this promise is a rather weak self-interested reason (e.g., the reason I have to avoid making an extra trip onto campus), and that the reason that I have to further my own self-interest is a non-moral reason. If moral reasons override non-moral reasons with respect to an act's moral status (that is, if moral reasons are morally overriding), then no modification of this situation in which only the strength and/or number of the relevant non-moral reasons are altered will change the fact that I am morally required to meet with the student. Thus, even if someone were to offer me a million dollars to give a lecture during the time at which I'd promised to meet with the student, I would still be morally required to meet with the student, foregoing this unique opportunity to make a million dollars. If moral reasons are morally overriding, then the non-moral reason that I have to make a million dollars is, no matter how strong, powerless to prevent the moral reason that I have to keep my promise from generating a moral requirement to do so, no matter how trivial this promise, and no matter how weak the moral reason stemming from it.<sup>32</sup>

<sup>31</sup> I will call any reason that is relevant to determining an act's deontic status a *morally relevant reason* but will reserve the label *moral reason* for only those reasons that, morally speaking, count for or against performing some action. This is an important distinction to make, because it may be that not all morally relevant reasons are moral reasons. That is, it may be that some reasons can justify performing acts that it would otherwise be morally impermissible to perform without themselves, morally speaking, counting in favor of (or against) those acts. For more on this issue, see Chapter 5.

<sup>32</sup> Since I want the reason that I have to make the million dollars to be a non-moral reason, let's assume that I must spend all of the money on personal luxuries such that only I will benefit.

MO is implausible, as I hope the above example illustrates. In any case, I'll be arguing against MO in Chapter 5. And it's important to keep in mind that I'm presuming that MO is false, for I readily admit that, were MO true, MR would be untenable. To see why, we need only note that MO and MR together entail the following implausible view:

**RO** If S has more moral reason to perform  $x$  than to perform  $y$  ( $y \neq x$ ), then S is not rationally permitted to perform  $y$ . And, thus, if S has most moral reason to perform  $x$ , then S is rationally required (i.e., has decisive reason, all things considered) to perform  $x$ .<sup>33</sup>

According to RO, moral reasons are *rationaly overriding* (hence, the abbreviation 'RO'). If this thesis were true, then it would always be *objectively irrational* (i.e., contrary to reason) to refrain from doing what one has most *moral* reason to do, even when what one has most moral reason to do is supported by only the most trivial of moral reasons and opposed by the weightiest of non-moral reasons.<sup>34</sup> This is implausible, for it seems that, in many instances, we have sufficient reason, all things considered, to act contrary to what we have most moral reason to do. It seems, for instance, that I would have sufficient, if not decisive, reason to break my promise to meet with a student to discuss his exam grade if this is the only way for me to take advantage of some unique opportunity to make a million dollars. So, given that RO is implausible, we had better not accept both MO and MR, which together entail RO. Thus, my case in favor of MR is contingent on our finding MO implausible. If what I've said so far isn't enough to convince you of the implausibility of MO, then, for now, you'll just have to accept my promise to provide more compelling reasons for rejecting MO in Chapter 5.

Some philosophers don't conflate MR with MO, but they do something that's nearly as bad: they suggest that MR is inextricably linked

<sup>33</sup> Actually, the first sentence in RO follows straightforwardly from MO and MR only if we assume that the following claim is implicit in MR: If S is not morally permitted to perform  $y$ , then S is not rationally permitted to perform  $y$ . But since this claim is equivalent to 'If S is morally required to perform  $\sim y$ , then S is rationally required (i.e., has decisive reason, all things considered) to perform  $\sim y$ ', which is just an instance of the claim that is explicit in MR, it's safe to say that RO is entailed by MO and MR.

<sup>34</sup> An act is *objectively irrational* if and only if the agent has decisive (objective) reasons not to perform it. Whenever I use the words 'rational' and 'irrational' without qualification, I mean to be talking about what is *objectively* rational or irrational.



to the view that moral requirements are themselves *reason providing*—that is, to:

**RP** The fact that S is morally required to perform  $x$  constitutes (and, thus, provides) an overriding (and, hence, decisive) reason for S to perform  $x$ .

Stephen Darwall, for instance, links MR to the view that “moral obligations always give agents conclusive reasons for acting that outweigh or take priority over any potentially competing considerations” (2006a, p. 26). And David Brink does the same in writing: “Moral requirements provide agents with overriding reasons for action; necessarily, it is on balance irrational to act contrary to moral requirements” (1997, p. 256).<sup>35</sup> But the two are not inextricably linked; one can accept that there is always decisive reason to do what one is morally required to do (that is, MR) while denying that moral requirements provide agents with overriding reasons for abiding by them (that is, RP). The fact that S is morally required to perform  $x$  may indicate that there are certain sorts of reasons for S to perform  $x$  that are indeed decisive, but the fact the S is morally required to perform  $x$  does not *itself* constitute an additional reason (let alone a decisive one) for S to perform  $x$ —at least, not in my opinion.<sup>36</sup>

So although some may be compelled to accept MR because they think that moral requirements generate overriding reasons to abide by them, others, like myself, may be driven to accept MR because they think that morality is limited in what it can require of us—that morality can require us to do only that which we have decisive reason to do, all things considered. The thought would be that although moral requirements don’t generate overriding reasons to abide by them, MR is, nevertheless, true, for non-moral reasons serve to constrain what morality can require of us in that they sometimes successfully counter our moral reasons, preventing them from generating moral requirements.<sup>37</sup> But although non-moral

<sup>35</sup> Thomas Scanlon endorses RP. He says, “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” (1998, p. 1).

<sup>36</sup> Of course, I would allow for the possibility that the fact that S is morally required to perform  $x$  could in conjunction with certain other facts constitute a reason for S to perform  $x$ , as where, say, S desires to do what she is morally required to do and the fact that her performing  $x$  would fulfill one of her desires constitutes a reason for her to perform  $x$ .

<sup>37</sup> Recall that to say that the reasons to  $\varphi$  successfully counter the reasons to  $\psi$  is to say that the reasons to  $\varphi$  prevent the reasons to  $\psi$  from being decisive by, say, equaling, outweighing, undermining, or silencing them. Another possibility is that the reasons to  $\varphi$

reasons, such as, self-interested reasons, serve to constrain what morality can require of us, it doesn't follow that self-interest reigns supreme. It's plausible to suppose that although sometimes self-interested reasons will prove decisive (all things considered) and will thereby make it permissible for us to act contrary to what we have most moral reason to do, other times moral reasons will prove decisive (all things considered) and will thereby generate a moral requirement for us to act contrary to our self-interest, as was the case above, where Anita had a decisive moral reason to refrain from murdering her uncle.

Interestingly, if we deny both MO and RO, then MR becomes not only tenable, but nearly unassailable, for if we admit that non-moral reasons can prevent moral reasons from generating moral requirements (and this is, as I argue in Chapter 5, what we should say if we deny both MO and RO), then what besides MR offers a more plausible account of when they do? If MR is true, then non-moral reasons will prevent moral reasons from generating moral requirements when they successfully counter them. And if the non-moral reasons in favor of performing a given act don't justify performing that act—making it morally permissible to do so—when they successfully counter all other reasons (moral and non-moral) for doing otherwise, then when?

So, as I noted earlier, part of my case for MR is that it coheres well with other claims that I defend in the book, such as  $\sim$ RO and  $\sim$ MO. But, as I also mentioned, there are other independent grounds for accepting MR. Below, I offer an independent argument for MR. The argument is not original to me, but I hope to state it more precisely than it has heretofore been stated. Moreover, I hope to buttress the argument by providing an original argument for its most controversial premise. Before I can state either argument, though, I need to explain some of the relevant terminology.

First, to say that S performs  $\sim x$  is to say that S performs the "act" of refraining from performing  $x$ . So if  $x$  is, for instance, the act of pushing some specific button at some particular time, then  $\sim x$  is the "act" of refraining from pushing that button at that time.

Second, to say that S has *decisive reason* to  $\varphi$  is to say that S's reasons are such as to make S objectively rationally required to  $\varphi$ , and to say that S has *sufficient reason* to  $\varphi$  is to say that S's reasons to  $\varphi$  are such as to make S objectively rationally permitted to  $\varphi$ . Thus, S has decisive reason to  $\varphi$  if

are incommensurable with the reasons to  $\psi$  such that there is sufficient reason both to  $\varphi$  and to  $\psi$ .

and only if S does not have sufficient reason to  $\sim\varphi$ . I'll use 'S has decisive reason to  $\varphi$ ' and 'S has sufficient reason to  $\varphi$ ' as shorthand for 'S has decisive reason to  $\varphi$ , all things considered' and 'S has sufficient reason to  $\varphi$ , all things considered', respectively.

Third, to say that S has sufficient reason to perform  $\sim x$  is to say that there is at least one act-token that is an instance of S's refraining from performing  $x$  that S has sufficient reason to perform. Likewise, to say that S has decisive reason to perform  $\sim x$  is to say that there is at least one act-token that is an instance of S's refraining from performing  $x$  that S has decisive reason to perform.

Fourth, to say that S *freely performs*  $x$  (or  $\sim x$ )<sup>38</sup> is to say that S performs  $x$  having the relevant sort of control over whether or not she performs  $x$ —that is, the sort of control that is necessary for her being an appropriate candidate for praise or blame with respect to her having performed  $x$ . More specifically, to say that S freely performs  $x$  is to say that S satisfies whatever other conditions in addition to having the relevant sort of knowledge (such as knowledge about what she is doing and about what she is bringing about) that is necessary for her to be morally responsible for having performed  $x$ .

Fifth, to say that S *knowledgeably performs*  $x$  is to say that S performs  $x$  knowing all the relevant facts—the relevant facts being those facts the ignorance of which would either inculpate or exculpate her for performing  $x$ .

Sixth, to say that S is *blameworthy* for performing  $x$  is to say both that it is appropriate for S to feel guilt about having performed  $x$  and that it is appropriate for others to feel indignation—and, perhaps, also resentment—in response to S's having performed  $x$ . And, here, I use 'appropriate' in the sense of being apt, fitting, or correct and, thus, in the same sense that fear is the appropriate response to the perception of danger. In this sense, it can be appropriate to blame oneself or someone else even though having this attitude (and/or expressing it) would not be instrumental in bringing about any good.<sup>39</sup>

<sup>38</sup> Hereafter, I'll leave the '(or  $\sim x$ )' implicit.

<sup>39</sup> It's important to keep distinct the issue of whether it is appropriate to blame someone for doing  $x$  and the issue of whether it is appropriate to intend to perform an act that constitutes the outward expression of this attitude. One can be appropriate without the other being appropriate. See §§3.2.3 for more on this matter. And see BENNETT 1979, STRAWSON 1962, and WALLACE 1994 for criticisms of the view that blame is appropriate if and only if the outward expression of this attitude would lead to a desired change in the agent and/or her behavior.

Seventh and last, when I say that S is morally required to perform  $x$ , I mean this in the objective sense—that is, I mean that S is *objectively* morally required to perform  $x$ .

With these terms defined, I can now state the argument as follows:

- 2.10 If S is morally required to perform  $x$ , then S would be blameworthy for freely and knowledgeably performing  $\sim x$ .<sup>40</sup>
- 2.11 S would be blameworthy for freely and knowledgeably  $\varphi$ -ing only if S does not have sufficient reason to  $\varphi$ .<sup>41</sup>
- 2.12 So, if S is morally required to perform  $x$ , then S does not have sufficient reason to perform  $\sim x$ . (From 2.10 and 2.11)
- 2.13 If S does not have sufficient reason to perform  $\sim x$ , then S has decisive reason to perform  $x$ . (From the definitions of ‘sufficient reason’ and ‘decisive reason’ above)
- 2.14 Therefore, if S is morally required to perform  $x$ , then S has decisive reason to perform  $x$ —and this is just MR. (From 2.12 and 2.13)<sup>42</sup>

The argument is deductively valid, and it seems that 2.11 is the only assumption that isn’t a conceptual truth. Premise 2.13 is clearly a conceptual truth. And 2.10 expresses the common assumption that there is a conceptual connection between blameworthiness and wrongdoing.<sup>43</sup> Although there may not be an essential connection between blameworthiness and wrongdoing *per se*, there is, it seems, an essential connection between blameworthiness and *freely and knowledgeably* doing what’s wrong. And if this is right, then anyone who denies MR is committed to denying 2.11. Yet 2.11 is intuitively plausible—even those who deny MR admit as much.<sup>44</sup> Besides, there is, I think, a sound argument

<sup>40</sup> ‘S would be blameworthy for freely and knowledgeably performing  $\sim x$ ’ should be read as shorthand for ‘S would be blameworthy if S were to freely and knowledgeably perform  $\sim x$ ’.

<sup>41</sup> Note that 2.11 entails: ‘S would be blameworthy for freely and knowledgeably performing  $\sim x$  only if S does not have sufficient reason to perform  $\sim x$ , all things considered’. One need only substitute  $\sim x$  for  $\varphi$ .

<sup>42</sup> This sort of argument for MR is not original to me. Darwall (2006a and 2006b, p. 292) and Skorupski (1999, pp. 170-171) make similar arguments for MR. And Shafer-Landau (2003, pp. 192-193) presents the same kind of argument but for a considerably weaker thesis: viz., if S is morally required to perform  $x$ , then S has a (not necessarily decisive or even sufficient) reason to perform  $x$ .

<sup>43</sup> See, for instance, DARWALL 2006a, GIBBARD 1990, MILL 1991 [1861], and SKORUPSKI 1999.

<sup>44</sup> For instance, David Sobel, who explicitly denies MR in D. SOBEL 2007, claims in D. SOBEL 2009 that it “seems quite intuitive that earnestly blaming a person for  $\varphi$ -ing entails the view

for 2.11, which is based on assumptions that are even more intuitively plausible than 2.11. But before I present that argument, let me rebuff some putative counterexamples to 2.11.<sup>45</sup>

Consider the following example. Suppose that Arthur, a white supremacist, sneaks up behind an unsuspecting black man, named Bert, and clubs him over the head, knocking him unconscious. He does so out of hatred for blacks. However, unbeknownst to Arthur, Bert was just about to shoot and kill his ex-girlfriend Carla, who, we'll suppose, is completely innocent. As it turns out, then, Arthur's act saves Carla's life. Arthur had, then, sufficient reason to club Bert over the head, for assume that any less violent or injurious act would have been insufficient to save Carla's life. Nevertheless, Arthur is clearly blameworthy. So this may seem to be a counterexample to 2.11. But note that Arthur doesn't *knowledgeably* club Bert over the head. Arthur is ignorant of the fact that his clubbing Bert over the head is necessary to save Carla's life, and it is his ignorance of this fact that inculpates him. The antecedent in 2.11 is, therefore, false, and so this is no counterexample to 2.11.

Of course, a better counterexample is certainly in the neighborhood. Imagine, for instance, a slightly different version of the case. In this version, Arthur knows all the relevant facts and, thus, freely *and knowledgeably* clubs Bert over the head. Nevertheless, what motivates him is not the thought that doing so will save Carla but only the desire to hurt a black person. Here, too, it seems that Arthur had sufficient reason to club

that the agent all things considered ought not to have  $\phi$ -ed." And he says, "It also seem quite intuitive to say that if one is acting as one has most reason to act, then one is acting as one ought and therefore one's action is not worthy of blame."

<sup>45</sup> Jussi Suikkanen and Paul McNamara have both suggested that 2.11 might have problematic implications in the following sort of case. Suppose that I accidentally (and, perhaps, also non-negligently) make two incompatible appointments: A and B. They're incompatible, because it's impossible for me to keep both them given that they're for the same time and in different locations. Assume, though, that A and B are equally important. Given their equal importance, it may seem that I have sufficient reason to break A as well as sufficient reason to break B. And, thus, according to 2.11, neither would be blameworthy. Yet, surely, this is a moral dilemma if ever there was one, and so I would be blameworthy no matter which appointment I break. So 2.11 seems to rule out moral dilemmas. But, in fact, as Howard Nye has pointed out to me, 2.11 rules out moral dilemmas only if there are no rational dilemmas, and it's unclear why anyone would want to claim that there can be moral dilemmas but no rational dilemmas. So if someone wants to claim that I would be blameworthy no matter which appointment I break, then that someone should also claim that I lack sufficient reason to break either appointment and so would be doing something contrary reason no matter which appointment I break.

Bert over the head, and yet clearly he is blameworthy.<sup>46</sup> But although it's clear that Arthur is blameworthy, it's far from clear that Arthur is blameworthy for clubbing Bert over the head. I think that we can rightly blame Arthur for his vicious motive, for his malevolent intent, and for his racist attitudes.<sup>47</sup> We can even rightly blame him for acting out of hatred and malice. But I don't think that we can rightly blame him for clubbing Bert over the head, as this is exactly what he should have done. Later in life, when Arthur finally comes to realize the error of his ways, what he should come to regret and feel guilty about is the fact that he was a racist who acted out of hatred and malice. Arthur should not, however, regret having clubbed Bert over the head, nor should he feel guilty for having done so, for this is what he had sufficient (indeed, decisive) reason to do. To accept blame for having  $\phi$ -ed, one must judge that one should not have  $\phi$ -ed. But although he should neither have wished Bert harm nor acted out of hatred or malice, he should have clubbed him over the head, as this was necessary to save Carla's life.

In arguing that Arthur should be blamed, not for clubbing Bert over his head, but for his malicious intent, his racist attitudes, and his acting out of malice, I've appealed to 2.11. This may seem illegitimate, but I'm trying to establish only that if one finds 2.11 intuitively plausible, the above examples shouldn't dissuade one from accepting 2.11. Although it's clear that Arthur is blameworthy for something, it's not clear that he is blameworthy for freely and knowledgeably doing anything that he had sufficient reason to do. Thus, the proponent of 2.11 can reasonably claim that there is no  $\phi$  such that Arthur had sufficient reason to  $\phi$  and yet is

<sup>46</sup> I thank Mike Almeida, Richard Chappell, and Clayton Littlejohn for proposing these sorts of putative counterexamples. Another type of putative counterexample, suggested to me by Peter de Marneffe, is Sophie's choice. It may seem both that Sophie had sufficient reason to choose to save her son and that it is, nevertheless, appropriate for her to feel guilty for having made this choice. This, however, is a tricky case. My suspicion is that although we may think her guilt appropriate insofar as we question whether she truly had sufficient reason to sacrifice her daughter, we shouldn't think it *appropriate* for her to feel guilt (in the sense of its being apt as opposed to psychologically normal) insofar we think that she did have sufficient reason to make the choice that she made. But it does seem reasonable to question whether she in fact had sufficient reason to make the choice that she made. We might think, for instance, that she had decisive reason to give each child an equal chance at being saved and that it is, therefore, appropriate for her to feel guilty insofar as she may have failed to have done so—perhaps, she chose to save her son because he was her favorite. Or we might think that a mother never has sufficient reason to sacrifice her child, not even for the sake of ensuring that at least one of her two children lives.

<sup>47</sup> For how we can be held morally responsible for our attitudes, see SCANLON 1998, A. M. SMITH 2005, and A. M. SMITH 2008.

blameworthy for freely and knowledgeably  $\varphi$ -ing. If, on the one hand, we let ' $\varphi$ ' stand for 'clubbing Bert over the head', then although it's clear that Arthur had sufficient reason to  $\varphi$ , it's not so clear that he is blameworthy for freely and knowledgeably  $\varphi$ -ing. And if, on the other hand, we let ' $\varphi$ ' stand for something such as 'acting out of malice', then although it's clear that Arthur is blameworthy for freely and knowledgeably  $\varphi$ -ing, it's not so clear that he had sufficient reason to  $\varphi$ .<sup>48</sup> So I don't see any clear counterexample to 2.11. Having cleared that up, let me explain why we should accept 2.11.

The thought underlying 2.11 is that agents are blameworthy only for freely and knowledgeably doing what they lack sufficient reason to do. One way to bring out the intuitive plausibility of this claim is to point to the tension there is in blaming someone for acting in a certain way while acknowledging that she had sufficient reason to act in that way. Stephen Darwall puts the point thusly:

It seems incoherent...to blame while allowing that the wrong action, although recommended against by some reasons, was nonetheless the sensible thing to do, all things considered.... Part of what one does in blaming is simply to say that the person shouldn't have done what he did, other reasons to the contrary notwithstanding. After all, if someone can show that he had good and sufficient reasons for acting as he did, it would seem that he *has* accounted for himself and defeated any claim that he is to blame for anything. Accepting blame involves an acknowledgment of this proposition also. To feel guilt is, in part, to feel that one shouldn't have done what one did. (2006b, p. 292)<sup>49</sup>

Another way to bring out the intuitive plausibility of 2.11 is to point to the tension there is in holding someone morally responsible for her actions on account of her having the capacity to respond appropriately to the relevant reasons and then blaming her for responding appropriately to the relevant reasons by doing what she had sufficient reason to do. Let me explain. It seems that an agent can be blameworthy for her actions only if she is morally responsible for them and that she can be morally responsible for them only if she has the relevant sort of control over her actions.<sup>50</sup>

<sup>48</sup> Note that saying that Arthur acted out of malice is not equivalent to saying that Arthur clubbed Bert over the head. These are not two equivalent descriptions of the same act, for Arthur can club Bert over the head without acting out of malice.

<sup>49</sup> See also DARWALL 2006a, especially p. 98.

<sup>50</sup> Some deny that control is a necessary condition for moral responsibility. Some argue that we can appropriately be held responsible for our judgment-sensitive attitudes even if we do not have volitional control over them—see, for instance, SCANLON 1998, A. M. SMITH 2005, and A. M. SMITH 2008. The thought is that an agent is morally responsible for something if

What's more, it's plausible to suppose that she has the relevant sort of control over her actions only if she has the capacity to respond appropriately to the relevant reasons for action—that is, the capacity both to recognize what the relevant reasons are and to react appropriately to them, being moved by them in accordance with their associated strengths.<sup>51</sup> Indeed, it's this capacity for responding appropriately to the relevant reasons, which normal adult humans typically possess and which children, animals, and the criminally insane typically lack, that distinguishes those who can potentially be blameworthy for their actions from those who can't.

Now, given that an agent can be blameworthy only if she has the capacity to respond appropriately to the relevant reasons, it's important to note that, in flawlessly exercising this capacity, she could be led to perform any act that she has sufficient reason to perform. Indeed, if there is but one act that she has sufficient reason to perform, then this capacity will, if exercised flawlessly, lead her both to recognize that this is the only act that she has sufficient reason to perform and to react by performing it. But if an agent is morally responsible and, thus, potentially blameworthy in virtue of having the capacity to respond appropriately to the relevant reasons, then how can we rightly blame her for doing something that she has sufficient reason to do when, in flawlessly exercising this very capacity, this is what she is led to do? Surely, it can't be that the very capacity that opens the door to an agent's being blameworthy is the one that leads her to perform blameworthy acts. And yet that's exactly what we would be

and only if that something is rightly attributable to her in that it expresses her judgments, values, or normative commitments. This view is known as attributionism, and it contrasts with volitionalism (see, e.g., LEVY 2005). But even attributionists allow that volitional control can be relevant in determining moral responsibility for bodily movements. Angel M. Smith, for instance, says: "In some cases (e.g., in determining a person's responsibility for a bodily movement), it may make sense to ask whether the agent has *voluntarily chosen* the thing in question, because that will determine whether that thing can reasonably be taken to express her judgments" (2008, p. 368). In any case, what matters to attributionists is whether the action is connected to the agent's underlying normative judgments such that she can, in principle, be called upon to defend it with reasons and be blamed if no adequate defense can be provided (A. M. SMITH 2008, p. 370). Given that citing sufficient reason for performing an action is presumably an adequate defense of one's performing that action, attributionists seem committed to my claim that S would be blameworthy for freely and knowledgeably  $\varphi$ -ing only if S does not have sufficient reason to  $\varphi$ —that is, to 2.11.

<sup>51</sup> We may need to talk about the capacities of mechanisms rather than the capacities of agents—see FISCHER AND RAVIZZA 1998, especially p. 38. If so, my argument could easily be revised so as to replace all talk of the agent's capacity to respond appropriately to the relevant reasons with talk of the reasons-responsiveness of the mechanism that issued in the agent's action.



allowing if we held that agents can be blameworthy for performing acts that they have sufficient reason to perform, for it is their capacity to respond appropriately to the relevant reasons that opens the door to the their potentially being blameworthy in the first place—a capacity that, when exercised flawlessly, leads them to perform acts that they have sufficient reason to perform.

To put the point in slightly different terms, it seems inappropriate to hold an agent responsible on the condition that she has the capacity to be guided by sound practical reasoning and then blame her for acting as she might very well be led to act if she is guided by sound practical reasoning. And since an agent can be led to perform any act that she has sufficient reason to perform when guided by sound practical reasoning, it seems inappropriate to blame her for freely and knowledgeably doing what she has sufficient reason to do.

More formally, the argument for 2.11 is this:

- 2.15 S is morally responsible for whether or not she  $\varphi$ s and, thus, potentially blameworthy for  $\varphi$ -ing or failing to  $\varphi$  only if S has the relevant sort of control over whether or not she  $\varphi$ s.
- 2.16 S has the relevant sort of control over whether or not she  $\varphi$ s only if S has the capacity to respond appropriately to the relevant reasons (both moral and non-moral).<sup>52</sup>
- 2.17 So, S is potentially blameworthy for whether or not she  $\varphi$ s only if S has the capacity to respond appropriately to the relevant reasons. (From 2.15 and 2.16)
- 2.18 If S has sufficient reason to  $\varphi$ , then, in flawlessly exercising her capacity to respond appropriately to the relevant reasons, S could be led to freely and knowledgeably  $\varphi$ .<sup>53</sup>
- 2.19 If S is potentially blameworthy for whether or not she  $\varphi$ s only if she has the capacity to respond appropriately to the relevant reasons, then S can't be blameworthy for freely and

<sup>52</sup> For a further defense of 2.15 and 2.16, see FISCHER AND RAVIZZA 1998. Note that it's not enough for an agent to have the capacity to respond appropriately to non-moral reasons. To be morally responsible, an agent must have the capacity to respond, and respond appropriately, to moral reasons as well. Certain psychopaths have the capacity to respond appropriately to non-moral reasons but are incapable of recognizing the fact that the rights and interests of others provide them with reasons for acting in ways that respect their rights and promote their interests. Such psychopaths are not morally responsible for their actions—see FISCHER AND RAVIZZA 1998, pp. 76-81.

<sup>53</sup> Whether she would or not depends on whether she also has sufficient reason to  $\varphi$ . It could be, after all, that she has sufficient reason both to  $\varphi$  and to refrain from  $\varphi$ -ing.

knowledgeably  $\varphi$ -ing when, in flawlessly exercising this capacity, S could be led to freely and knowledgeably  $\varphi$ .

**2.20** So, S wouldn't be blameworthy for freely and knowledgeably  $\varphi$ -ing if S has sufficient reason to  $\varphi$ . (From 2.17, 2.18, and 2.19<sup>54</sup>)

**2.21** Therefore, S would be blameworthy for freely and knowledgeably  $\varphi$ -ing only if S does not have sufficient reason to  $\varphi$ —and this is just 2.11. (From 2.20 by contraposition)

In considering this argument, it will be helpful to have a specific example in mind. So consider a revised version of the example from above, where I've promised to meet with a student to discuss his exam grade but have the unique opportunity to make some sum of money by giving a lecture instead. Assume that I'm offered just enough money so that I have just as much reason, all things considered, to give the lecture as to show up for my meeting with the student as promised. And assume that I have considerably less reason to do anything else. That is, assume that I have both sufficient reason to keep my promise and sufficient reason to give the lecture and insufficient reason to do anything else. Furthermore, assume that I'm strongly responsive to reasons, meaning both that I always recognize what the relevant reasons are and that I'm always moved to act in accordance with their associated strengths, such that I always do what I have decisive reason to do, only do what I have sufficient reason to do, and choose arbitrarily which of two acts to perform (as by, say, the toss of a coin) if and only if I have sufficient and equivalent reason to do either. Thus, being strongly reasons-responsive and having sufficient and equivalent reason to do either, I choose arbitrarily to give the lecture. That is, I first designate heads for giving the lecture and designate tails for keeping my promise and then toss a coin, which lands heads. As a result, I decide to give the lecture.

Of course, being strongly reasons-responsive I would have kept my promise and met with the student had there been decisive reason to do so. For instance, I would have kept my promise had my moral reason for doing so been a bit stronger, as where, say, the student's scholarship was at stake. Likewise, I would have kept my promise had my self-interested reason for giving the lecture been a bit weaker, as where, say, the monetary offer had been a bit less. And had there been a third option that I had decisive reason to do (e.g., saving some drowning child), I would have

<sup>54</sup> Roughly, the structure of the argument from 2.17-2.19 to 2.20 goes something like this:  $B \rightarrow C, S \rightarrow L, (B \rightarrow C) \rightarrow (L \rightarrow \sim B) \therefore S \rightarrow \sim B$ .

done that instead. But, as it was, none of these were the case. As it was, I had sufficient reason to break my promise and give the lecture. And that's what I did, having flawlessly exercised my capacity for sound practical reasoning.

Given my flawless execution of my capacity for sound practical reasoning, how can I be faulted for breaking my promise? Isn't it inappropriate to hold me morally responsible and, thus, potentially blameworthy in virtue of my capacity for being guided by sound practical reasoning and then blame me for acting as sound practical reasoning leads me to act? Shouldn't even the student admit that, had he been in my situation and perfectly rational, he might also have been led to act as I did? And if so, how can he rightfully resent me for acting as he would have acted?

Of course, things would have been different had there been decisive reason for me to have kept my promise. In that case, it would have been appropriate for the student to resent me for breaking my promise to him. If I had decisive reason to keep my promise as well as the capacity to respond appropriately by recognizing this fact and reacting appropriately to it, then I must have failed to have done something that I could, and should, have done. Perhaps, I failed to recognize the fact that I had decisive reason to keep my promise even though I was capable of doing so. Or, perhaps, I recognized that I had decisive reason to do so, but failed to react appropriately even though I had the capacity to do so—perhaps, I was weak-willed and gave into the temptation to earn some easy money. In either case, I would have failed in some way and could, then, rightfully be blamed for this failure. So I see how it can be appropriate to blame someone for failing to do what she had decisive reason to do, but I don't see how we can rightfully blame someone for freely and knowledgeably doing what she had sufficient reason to do.

Admittedly, lots of people perform acts that they have sufficient reason to perform and are, nonetheless, blameworthy, because they were motivated by some consideration that shouldn't have motivated them. Such is the case in the example from above where Arthur does what he has sufficient reason to do out of a desire to see people of a certain race harmed. But, as I noted above, I think that in this case we should not blame Arthur for doing anything that he had sufficient reason to do (such as, clubbing Bert over the head). The alternative would be to allow that people can be blameworthy for freely and knowledgeably doing what they had sufficient reason to do, and that, I've argued, is inappropriate insofar as the very capacity that makes them subject to blame can lead them, when

exercised flawlessly, to do what they have sufficient reason to do. The only way to forestall such a possibility is to insist upon MR, which, I've argued, we should.

Still, the person who wishes to deny MR might balk at 2.16, arguing that when it comes to being morally responsible, what's relevant is not the capacity to respond in a *rationally* appropriate manner, but rather only the capacity to respond in a *morally* appropriate manner.<sup>55</sup> Of course, this person must insist that the two can come apart and that, although doing what one recognizes as one's having sufficient reason to do is always a rationally appropriate response, it is sometimes a morally inappropriate response.<sup>56</sup> But why think this? Why think that it is morally inappropriate to respond to moral reasons as it is rationally appropriate to respond to them? A moral reason is just a certain type of reason, and, as a reason, it has a certain weight. So why think that it is morally inappropriate to give moral reasons only as much weight as they in fact have?

In any case, when we talk of an agent's needing control over her actions in order to be held morally responsible for them, the relevant sort of control seems to be *rational* self-control—that is, the capacity to respond in a rationally appropriate manner to the relevant reasons. If it were

<sup>55</sup> To be clear, 'to respond appropriately' should be taken as shorthand for 'to respond in a rationally appropriate manner' both in 2.16 and throughout the rest of the argument.

<sup>56</sup> I thank Pat Greenspan for raising this possible objection to 2.16. As far as I know, however, no one has defended such a position in print. Even those who think that being morally (or normatively) competent is a necessary condition for being morally responsible stop well short of suggesting that this competency requires the capacity to respond by giving moral reasons *more* than their due rational weight. Rather, they suggest, to the contrary, that agents must have the ability to grasp and respond to both moral *and* non-moral reasons—see, for instance, FISCHER AND RAVIZZA 1998 (chap. 3) and WOLF 1990 (chap. 6). Indeed, at least one author—viz., Susan Wolf—explicitly claims that responsibility entails the ability to appreciate and act in accordance with *all* the reasons there are, including both moral and non-moral reasons—see WOLF 1990, p. 141. On Wolf's Reason View of responsibility, "*full* freedom and responsibility will involve the ability to appreciate reasons that come from a variety of sources" (1990, p. 141) and she explicitly notes that "our image of the agent who is most able to see and understand what reasons there are need not coincide with that of the agent who is most acutely sensitive particularly to *moral* reasons" (1990, p. 137). And although she claims that freedom and responsibility involves the ability to appreciate and act in accordance with what she calls the "*True and the Good*," she makes clear that "[a]ppreciation of the Good need not be confined to appreciation of the *moral* good. Indeed, in certain contexts, appreciation of the moral good may interfere with one's ability to appreciate the nonmoral good or with one's ability to recognize reasons for preferring a morally inferior course of action" (1990, p. 137). Thus, being overly sensitive to moral reasons can diminish one's freedom and responsibility by diminishing one's rational self-control.

anything else, then someone could be strongly reasons-responsive and, for that reason, lack the sort of control that's essential for moral responsibility. But the idea that an agent could fail to be morally responsible on account of her being strongly reasons-responsive is quite counterintuitive.

## §2.6 Some concluding remarks

I've claimed that there is an essential connection between blameworthiness and freely and knowledgeably violating a moral requirement (see 2.10). And I've argued that given this essential connection and the fact that it is inappropriate to blame agents for freely and knowledgeably doing what they have sufficient reason to do (see 2.11), we should accept MR. In closing, I want to do three things. First, I want to explain why we can't similarly argue for the legal analogue of MR. Second, I want to show how the truth of MR and the falsity of this legal analogue allows us to explain why moral requirements, but not legal requirements, can be defeated by the presence of sufficient reasons for the agent's doing otherwise. And, third, I want to explain why all would not be lost even if MR turns out to be false. I'll take each in turn.

The legal analogue of MR is *legal rationalism*:

**LR** If S is legally required to perform  $x$ , then S has decisive reason to perform  $x$ , all things considered.

The reason why MR is plausible, whereas LR is implausible, is that whereas there is an essential connection between freely and knowledgeably violating a *moral* requirement and the appropriateness of emotions such as guilt, resentment, and indignation, there is no essential connection between freely and knowledgeably violating a *legal* requirement and the appropriateness of such reactive attitudes.

An agent can freely and knowledgeably violate a legal requirement without its being appropriate either for her to feel guilt or for others to feel indignation. To illustrate, consider that if in my rush to get my critically injured daughter to the emergency room I proceed through an intersection against the red light after looking carefully to ensure that I can do so safely, I will have freely and knowledgeably broken the law. But I should not feel guilty for having done so. Nor should others be indignant with me for having done so. After all, I did not put anyone at risk, as I did slow down and look carefully before proceeding against the red light. And although I did break the law, it can't be claimed that in so doing I failed to show

adequate respect for the law, for we're assuming that I had decisive reason to break the law in this instance.

So, it should come as no surprise that we can argue for MR on the assumption that we can't rightly be blamed for freely and knowledgeably doing what we have sufficient reason to do, but that we can't similarly argue for LR given the lack of any essential connection between the violation of a legal requirement and the appropriateness of such reactive attitudes.

Interestingly, the truth of MR along with the falsity of LR allows us to explain why moral requirements, but not legal requirements, can be defeated by the presence of sufficient reasons for the agent's doing otherwise. To illustrate how moral requirements can be defeated in this way, consider the following example. It is Tuesday morning and Professor Collins is lamenting the fact that he has to drive to campus this afternoon to hold office hours, when he gets a call from a friend informing him of a rare opportunity to meet Hall-of-Famer Reggie Jackson. Now Collins is a huge fan who has for many years wanted to meet Jackson. Meeting him would be one of the highlights of his life. Unfortunately, it's too late to cancel his office hours. So, if he doesn't hold them, some students may show up and wait in vain.<sup>57</sup>

So what should he do? Should he go and meet his childhood hero, or should he instead go to his office just in case some of his students want to complain about their grades? Intuitively, it seems that, in light of how important it is to him, he has sufficient reason, all things considered, to take advantage of this once-in-a-lifetime opportunity to meet his childhood hero. What's more, it seems morally permissible for him to do so as well, and this seems true even if we suppose that his going to meet his childhood hero will benefit no one besides himself. Interestingly, though, we think that, if he didn't have sufficient reason to be somewhere else, he would be morally required to hold his office hours. After all, he has at least a *prima facie* obligation to fulfill his commitments and agreeing to hold office hours is a kind of commitment. So why, in this case, doesn't the moral reason that he has to fulfill his commitment to hold office hours generate a moral requirement for him to do so?

The explanation can't be that Collins has as much moral reason to meet Jackson as to fulfill his commitment to hold office hours, for it seems that the self-interested reason that he has to meet Jackson is not a moral reason at all. Consider that if he were choosing between meeting Jackson

<sup>57</sup> This case comes from PORTMORE 2003 and is adapted from WOLF 1986, p. 142.

and staying home and staring at the ceiling, he wouldn't be morally obligated to meet Jackson. Nor would his meeting Jackson be morally supererogatory. Indeed, it seems that the fact that he would personally benefit from meeting Jackson counts not at all in favor of his meeting Jackson—not morally speaking, that is. Thus, it seems that his reason to meet Jackson is a non-moral reason.

The explanation, then, for why Collins isn't morally required to hold office hours seems to be that the self-interested and non-moral reason that he has to meet Jackson, being sufficiently strong, morally justifies his failing to fulfill his commitment to hold office hours. Thus, we seem to be implicitly committed to MR—that is, to the idea that an agent can't be morally required to do what she has sufficient reason not to do.

If instead MR were false, we would expect that if we were to take a case in which an agent has decisive reason to fulfill a moral requirement and then imagine a series of variants on that case in which the agent has ever stronger non-moral reasons to do something else, we would eventually arrive at a case in which the agent had sufficient reason to violate that moral requirement. But we never do arrive at such a case. For as soon as we are willing to say that the agent has sufficient reason to do that something else, we are no longer willing to say that there is a moral requirement to refrain from doing that something else.

To illustrate, consider the case in which I have promised to hold office hours and have nothing better to do. Clearly, in this case, what I should do, all things considered, is hold my office hours as promised, and it seems that I am morally required to do so as well. But now imagine a series of variants on this case in which I have an ever stronger self-interested reason to do something else. At some point, we are willing to say that I now have sufficient reason to violate my commitment to hold office hours. Yet once we get to this point, we are no longer willing to say that I am morally required to hold these office hours. The Reggie Jackson example is a case in point. The self-interested and non-moral reason that Collins has to take advantage of this once-in-a-lifetime opportunity to meet his childhood hero successfully counters the moral reason that he has to hold office hours, the result being that he is neither rationally nor morally required to hold office hours. (Of course, this is not to deny that he would be morally required to make it up to those students who showed up and waited in vain.) It seems, then, that an agent's having sufficient reason to refrain from performing some act defeats what would otherwise be a moral requirement for her to perform that act.

Compare this to what effect an agent's having sufficient reason to do otherwise has on legal requirements. As we did above, let's start with a case in which I have decisive reason to obey the requirement—in this case, a legal requirement. Imagine, then, a case in which I come upon a red light while not in any hurry and know that, if were to proceed against the red light, I would certainly get a ticket, as there are red-light cameras installed at the intersection. Now, imagine a series of variants on this case in which I have ever stronger and eventually sufficient reason to proceed against the red light. At the point at which we think there is sufficient reason for me to proceed against the red light, we don't conclude that there is no longer any legal requirement for me to wait for the green to proceed. In this respect, legal requirements differ from moral requirements. Legal requirements are not defeated by the agent's having sufficient reason to do otherwise. But what accounts for the difference? It seems to be the fact that whereas MR is true, LR is false. So besides the fact that MR is supported by the argument from blameworthiness, we should accept MR because it helps us to explain why non-moral reasons can, when sufficiently strong, prevent moral reasons from generating moral requirements.<sup>58</sup>

But what if I'm wrong about all of this? What if in spite of all my arguments to the contrary MR turns out to be false? If MR is false, then we don't have decisive reason to abide by all of morality's requirements. And, in that case, it seems to me that what we need to know is which moral requirements are the ones that we have decisive reason to obey, for it is only those that we just plain ought to obey. In that case, let's call the moral requirements that agents have decisive reason to obey *schmoral requirements*. I think that if I were convinced that MR was false, I would be more interested in giving an account of *schmorality* than in giving an account of morality. After all, it would seem, then, that providing a theory of *schmorality* is of much greater practical importance than providing an account of morality. And, in giving an account of *schmorality*, I could certainly appeal to *schmoral rationalism*, which would be analytically true:

**SR** If S is schmorally required to perform  $x$ , then S has decisive reason to perform  $x$ , all things considered.

So if MR is false, then I suggest replacing 'schmoral' for 'moral' and 'SR' for 'MR' throughout the book. I don't think that anything of practical significance would be lost in doing so.

<sup>58</sup> I'll have more to say on this topic in Chapter 5.



At this point, though, one might wonder why we shouldn't just focus on what we have decisive reason to do and pay no mind at all to whether or not it would be morally wrong to refrain from doing what we have decisive reason to do. In response, I would argue that, as rational beings, we're not only interested in what we ought to *do*, but also in how we ought to *feel* and, in particular, whether we ought to feel guilt, indignation, and/or resentment in response to someone's failing to do what she had decisive reason to do. If we were interested only in what we ought to do, then we could just ignore morality and focus solely on what we have decisive reason to do. But even if not particularly relevant to whether we ought to do something, the fact that an act is morally (or schmorally) required is highly relevant to the appropriateness of feeling various blaming emotions, such as guilt, indignation, and resentment.<sup>59</sup> Someone can freely and knowledgeably do something contrary to what she has decisive reason to do, but unless that act is also wrong (or schmorally impermissible) it would be inappropriate to blame her for acting as she did. After all, not every act that is contrary to reason is also wrong. For instance, many imprudent acts are, it seems, contrary to what there is decisive reason to do but are not morally wrong. So even if in wondering what we ought to do we can just ignore morality, we can't ignore morality if we want to know when it's appropriate to feel various blaming emotions, such as guilt, indignation, and resentment.



In this chapter, I've argued both that we should accept MR and that all traditional versions of consequentialism require us to perform acts that we lack decisive reason to perform. Therefore, we should reject all traditional versions of consequentialism. But I've also argued that MR compels us to accept consequentialism—at least, it does if we accept a certain plausible conception of practical reasons. In the next chapter, I'll examine this conception of practical reasons more closely, explaining why it's attractive, why recent attacks on it have failed, and why we should think that it is uniquely positioned to offer the most unified and systemic account of our considered convictions about practical reasons.

<sup>59</sup> I think that if we were to reject MR, we would have to reject 2.10 as well and hold that when someone freely and knowledgeably acts wrongly, we can blame them only if they didn't do so in virtue of having sufficient reason to do so. Thus, if we reject MR, it is not moral requirements but schmoral requirements that would be conceptually tied to blame.

## CHAPTER 3

---

# The Teleological Conception of Practical Reasons

- 3.1 Getting clear on what the view is
- 3.2 Clearing up some misconceptions about the view
- 3.3 Scanlon's putative counterexamples to the view
- 3.4 Arguments for the view

IT IS THROUGH our actions that we affect the way the world goes. Indeed, whenever we face a choice of what to do, we also face a choice of which of various possible worlds to actualize.<sup>1</sup> Moreover, whenever we act *intentionally*, we act with the aim of making the world go a certain way. The aim needn't be anything having to do with the causal consequences of the act. The aim could be nothing more than to bring it about that one performs the act. For instance, one could intend to run merely for the sake of bringing it about that one runs. The fact remains, though, that for every intentional action there is some end at which the agent aims.

If our actions are the means by which we affect the way the world goes, and if our intentional actions necessarily aim at making the world go a certain way, then it is only natural to suppose that what we have most reason to do is determined by which is the way we have most reason to want the world to go. To put things more precisely, an agent's reasons for action are a function of her reasons for preferring certain possible worlds to others, such that what she has most reason to do is to bring about the possible world, which of all those available to her, is the one that she has

<sup>1</sup> I will assume that for each act available to an agent there is some determinate fact as to what the world would be like were she to perform that act. This assumption is sometimes called *counterfactual determinism*—see, e.g., BYKVIST 2003, p. 30. Although this assumption is controversial, nothing that I will say here hangs on it. I make the assumption only for the sake of simplifying the presentation. If counterfactual determinism is false, then instead of correlating each act with a unique possible world, we will need to correlate each act with a probability distribution over the set of possible worlds that might be actualized if S were to perform that act.

most reason to want to be actual. This is what's known as the *teleological conception of practical reasons*.<sup>2</sup>

Whether this is the correct conception of practical reasons is important not only in its own right, but also in virtue of its potential implications for moral theorizing. As I argued in the previous chapter, this conception of practical reasons in conjunction with moral rationalism (MR) compels us to accept consequentialism. In this chapter, I'll argue that we should accept the teleological conception of (practical) reasons—or 'TCR' for short. These arguments for TCR in conjunction with my arguments for MR in the previous chapter jointly constitute my argument for consequentialism.

TCR is, on its face, quite plausible. Even its critics admit as much.<sup>3</sup> Since TCR is *prima facie* plausible, my main task will be to show that it is not subject to the sorts of objections that critics have leveled against it. Thus, I'll spend most of this chapter rebutting objections and clearing up misconceptions. Nevertheless, I will, in the second half of the chapter, try to offer some positive arguments for it.

The chapter has the following structure. In §3.1, I offer a more precise statement of TCR, showing that the view consists in three distinct claims. And I explain why my formulation differs from those of some of its critics in eschewing talk of both value and desirability. Then, in §3.2, I clear up some common and potential misconceptions about the view. In §3.3, I rebut Scanlon's putative counterexamples to TCR, cases where putatively "many of the reasons bearing on an action concern not the desirability of outcomes but rather the eligibility or ineligibility of various other reasons" (SCANLON 1998, p. 84). And, finally, in §3.4, I provide arguments for each of TCR's three claims and for TCR as a whole.

### §3.1 Getting clear on what the view is

Let me start by stating the view as precisely as I can. Let  $a_1, a_2, \dots, a_n$  be the set of mutually exclusive and jointly exhaustive act alternatives available to a subject,  $S$ . Let  $o_1, o_2, \dots, o_n$  be their corresponding outcomes, where an act's outcome is construed broadly as the possible world that would be

<sup>2</sup> Whenever I speak of practical reasons or reasons for action, I will be referring to normative reasons for action. A normative reason for action is some fact that counts in favor of the agent's performing that action. Normative reasons contrast with explanatory reasons (i.e., facts that explain why the agent performed the action). One particularly important subclass of explanatory reasons is the set of motivating reasons, the facts that motivated the agent to perform the action—that is, the facts that the agent took to be her reasons for performing the action. See DARWALL 2006, p. 285.

<sup>3</sup> Scanlon, for instance, admits that this view "sounds plausible" (1998, p. 84).

actual were the act to be performed.<sup>4</sup> More precisely, then, the teleological conception of reasons can be stated as follows:

**TCR** (1) S has more reason to perform  $a_i$  than to perform  $a_j$  if S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains; (2) S has more reason to perform  $a_i$  than to perform  $a_j$  only if S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains; and (3) if S has more reason to perform  $a_i$  than to perform  $a_j$ , then this is so in virtue of the fact that S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains.<sup>5</sup>

More concisely, then, TCR is the view that S has more reason to perform  $a_i$  than to perform  $a_j$  if and only if, and because, S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains. For my purposes, though, it will be useful to keep the three claims separate; I'll refer to them as TCR-1, TCR-2, and TCR-3, respectively. And although not stated above, I take TCR to include the claim that S has a reason to perform  $a_i$  if and only if, and because, S has a reason to desire that  $o_i$ .

Having stated the view as precisely as I can, I will now proceed to clarify it, explaining in the process how and why it differs from T. M. Scanlon's formulation of the view, which goes as follows: "the purely teleological conception of reasons [is the view] according to which, since any rational action must aim at some result, reasons that bear on whether to perform an action must appeal to the desirability or undesirability of having that result occur, taking into account also the intrinsic value of the act itself" (SCANLON 1998, p. 84). Although my formulation differs from Scanlon's in how it's worded, I don't believe that it differs in substance from the view that he intended to describe—or so I will argue. My formulation differs from Scanlon's only in its degree of clarity and precision. There are, in fact, five separate points that need clarifying.

§§3.1.1 *Reasons for desiring as opposed to desirability*: Unlike Scanlon, I have formulated TCR in terms of the agent's reasons for desiring various possible outcomes as opposed to the desirability of those outcomes. To see why, we must first get clear on what the difference is. Let's start with what it means to say that an outcome is desirable. To say that that an outcome,

<sup>4</sup> Again, I'm assuming counterfactual determinism—see note 1 above.

<sup>5</sup> Michael Smith's (2006) formulation of the view is similar to mine. He says, "(x)(t)(x at t has all things considered reason to  $\varphi$  in circumstances C iff  $\varphi$ -ing is the unique action of those x can perform at t that brings about what x would desire most happens in C if he had a desire set that satisfied all requirements of...reason."

$o_i$ , is desirable is to say that it is fitting to desire that  $o_i$  obtains. And to say that it is fitting to desire that  $o_i$  obtains is just to say that there are sufficiently weighty reasons of the right kind to desire that  $o_i$  obtains.<sup>6</sup> What are the right kinds of reasons? They are all and only those reasons that are relevant to determining whether, and to what extent,  $o_i$  is desirable. Let's call these *fittingness reasons*.

I won't attempt to give a complete account of what sorts of reasons are, and what sorts of reasons are not, fittingness reasons. These are controversial issues, and I have nothing new to add to the growing debate. Even so, I can plausibly claim that there are some clear cases of what wouldn't count as fittingness reasons. First, if facts such as the fact that an evil demon will torture me unless I desire that  $o_i$  obtains constitute genuine reasons for me to desire that  $o_i$  obtains, then these are clearly not fittingness reasons, for such pragmatic "reasons" for desiring that  $o_i$  obtains clearly have no bearing on whether, or to what extent,  $o_i$ 's obtaining is desirable.<sup>7</sup> Now, as a matter of fact, I don't think that such pragmatic "reasons" do constitute genuine reasons for desiring that  $o_i$  obtains. I think instead that they constitute only reasons to want, and to act so as to cause oneself, to desire that  $o_i$  obtains. If, however, I'm wrong about this, then admittedly I'll need to revise TCR so as to exclude such pragmatic "reasons," for such reasons are no more relevant to whether one should act so as to bring it about that  $o_i$  obtains than they are to whether  $o_i$ 's obtaining is desirable.<sup>8</sup>

<sup>6</sup> More precisely, we should first say that, for all states of affairs  $p$  and  $q$ , it is better (i.e., preferable) that  $p$  is the case than that  $q$  is the case if and only if the set of all the *fittingness reasons* for preferring its being the case that  $p$  to its being the case that  $q$  is weightier than the set of all the *fittingness reasons* for preferring its being the case that  $q$  to its being the case that  $p$ . Then we can say that it is good (i.e., desirable) that  $p$  is the case if and only if the state of affairs in which it is the case that  $p$  is better than (i.e., preferable to) most of the states of affairs in the relevant contextually supplied comparison class. For more on this, see SCHROEDER 2008a and SCHROEDER 2008b.

<sup>7</sup> This is known as the *wrong-kind-of-reasons problem* for the fitting-attitude or buck-passing account of value (or desirability). For more on this problem and for some potential solutions to it, see RABINOWICZ AND RØNNOW-RASMUSSEN 2004, PARFIT 2008, and SCHROEDER 2008a.

<sup>8</sup> The revised version of TCR would, then, read as follows: S has more reason to perform  $a_i$  than to perform  $a_j$  if and only if, and because, the set of all the *non-pragmatic reasons* that S has for desiring that  $o_i$  obtains is weightier than the set of all the *non-pragmatic reasons* that S has for desiring that  $o_j$  obtains. Pragmatic reasons for S to desire that  $o_i$  obtains are reasons that are provided by facts about the consequences of S's desiring that  $o_i$  obtains, and non-pragmatic reasons are just reasons that are not pragmatic—for a more careful account of the relevant distinction, see STRATTON-LAKE 2005. The distinction that I'm drawing between pragmatic and non-pragmatic reasons is closely related to Derek Parfit's (2001) distinction

Second, it seems clear that agent-relative reasons are not fittingness reasons. To see why, consider that, in contrast to some stranger, I might have weightier (agent-relative) reasons to prefer the outcome in which my child lives to the outcome in which her child lives. But it would be odd to say that this is because the outcome in which my child goes on living is, other things being equal, better or more desirable than the outcome in which her child goes on living. Other things being equal, the outcome in which my child lives is neither more nor less desirable than the outcome in which her child lives. So although one can have agent-relative reasons for preferring one outcome to another, this does not entail that the one is better than, or preferable to, the other.<sup>9</sup> And, thus, agent-relative reasons for preferring one outcome to another are not fittingness reasons for preferring the one to the other.<sup>10</sup>

between state-given and object-given reasons as well as to Christian Piller's (2006) distinction between attitude-related and content-related reasons; it's not clear to me, though, that either is extensionally equivalent to mine.

As will be evident shortly, it's important to note that many agent-relative reasons for S to desire that  $o_i$  obtains, such as the fact that S's child will live if and only if  $o_i$  obtains, are non-pragmatic reasons and thus will not be excluded by the restriction to non-pragmatic reasons in the above revised version of TCR.

<sup>9</sup> This is known as the *partiality challenge* to the fitting-attitude or buck-passing account of value. See OLSON 2008, SUIKKANEN 2008, and ZIMMERMAN 2008 for some potential solutions to this particular problem.

<sup>10</sup> There are also time-relative reasons for preferring one outcome to another (and these too are not fittingness reasons). Sometimes, for instance, the preference that an agent should have before choosing to perform some action is the opposite of what it should be at some point after performing that action. Suppose, for instance, that Ana decided not to have an abortion even after learning that her fetus might have Down syndrome. Consequently, she gave birth to a boy, named Bill, with Down syndrome, who is now eight. She ought, at this point, be glad that she didn't have an abortion. After all, she does, at this point, have a very special bond with her son Bill, whom she has come to love for exactly who he is, which includes the fact that he has Down syndrome. If she could somehow take it all back and make the decision over, she should not wish/prefer (at this point) that she had had the abortion and then later given birth to a child without Down syndrome. Interestingly, though, this is compatible with our thinking that at the time of her initial decision, a time before she had formed any special relationship with what was then just an early-stage fetus, she should have preferred the outcome in which she had had the abortion and then given birth to a child without Down syndrome. That would have been the better outcome, and, at the time of her initial decision, she had no reason to prefer the worse outcome in which she has a child with Down syndrome. Given that what it is reasonable to prefer can change over time, TCR should actually be time-indexed. Consequently, TCR should be formulated as follows: S has more reason to perform  $a_i$  at  $t_i$  than to perform  $a_j$  at  $t_i$  if and only if, and because, S has more reason, at  $t_i$ , to desire that  $o_i$  obtains than to desire that  $o_j$  obtains. For more on the issue of how it can be reasonable to prefer that a loved one exists even though

It is this last exclusion that makes trouble for formulating TCR in terms of the value or the desirability of outcomes. If TCR is to be formulated in terms of value/desirability and value/desirability is to be understood exclusively in terms of agent-neutral reasons for desiring, then TCR will automatically disallow agent-relative reasons for action, such as the agent-relative reason that I have to save my own child as opposed to some stranger's child. Yet there is no reason why the teleologist should exclude the possibility of agent-relative reasons for action, as even the critics of TCR admit. Scanlon, for instance, says, "The teleological structure I have described is often taken to characterize not only 'the good' impartially understood, but also the good from a particular individual's point of view (the way she has reason to want things to go)" (1998, p. 81). So if we are to allow that one agent might have a reason to bring about a state of affairs that another has no reason to bring about and that one agent might have more reason to bring about some state of affairs than another agent does, we should eschew talk of the value/desirability of states of affairs and talk instead of the reasons that various agents have to desire those states of affairs, where some of these reasons will be agent-relative reasons.

Interestingly, Scanlon is aware that agent-relative reasons for valuing/desiring are not fittingness reasons. He says, "To claim that something is *valuable* (or that it is 'of value') is to claim that others also have reason to value it, as you do" (1998, 95). Furthermore, he claims that "[w]e can, quite properly, value some things more than others without claiming that they are more valuable" (1998, 95). We can properly value/desire some outcomes more than others without claiming that they are better than, or preferable to, those others. This is because which is most preferable (i.e., best) is a function of only our agent-neutral reasons for preferring them to the others, whereas which we have most reason to prefer is a function of both our agent-relative and our agent-neutral reasons for preferring them to the others.

As Jussi Suikkanen (2009, p. 6) argues, it is better (impersonally speaking) that  $o_i$  obtains than that  $o_j$  obtains if and only if it would be fitting for an impartial spectator to prefer  $o_i$  to  $o_j$ . An impartial spectator must be both impartial and a mere spectator. To ensure that she is impartial, we must assume that she has no personal relations with anyone involved. And to ensure that she is a mere spectator, we must assume that

one recognizes that there is a preferable state of affairs in which that loved one doesn't exist, see HARMAN 2009.

she is not involved either in bringing it about that  $o_i$  obtains or in bringing about that  $o_j$  obtains. Thus, the impartial spectator can have nothing but agent-neutral reasons for preferring one outcome to the other.

But, unlike the impartial spectator, situated agents can have agent-relative reasons for preferring one outcome to another given both their agential relations to those outcomes and their personal relations with those who would be better or worse off were those outcomes to obtain. So whereas it can, for instance, be fitting for me to prefer the outcome in which my child lives to the one in which some stranger's child lives given my personal relations with my child, it would not be fitting—at least, not if other things are equal—for an impartial spectator to have the same preference. And thus it can be proper for me to prefer the outcome in which my child lives even if this outcome is not better than (or preferable to) the one in which the stranger's child lives.

Now, if we can properly value/desire some outcomes more than others without claiming that they are better than, or preferable to, those others, then we should ask, "Why does Scanlon formulate the teleological conception of reasons in terms of value/desirability when he clearly wants to allow that the teleologist can accommodate agent-relative reasons for valuing/desiring and, consequently, agent-relative reasons for action?" The answer is that when Scanlon talks about value/desirability in the context of TCR, he means for this to include agent-relative value or what he refers to as "the good from a particular individual's point of view." Indeed, he brings up the teleological conception of reasons to explain why some think that we must assign agent-relative disvalue to an agent's killing in order to make sense of agent-centered restrictions against killing (1998, 94-95). But since Scanlon equates what is "good [or desirable] from a particular individual's point of view" with "the way she has reason to want things to go" (1998, p. 81), he should have no objection to my formulating TCR in terms of the agent's reasons for desiring. Indeed, given what he says, we should think that my formulating TCR in terms of the agent's reasons for desiring various outcomes is equivalent to his formulating TCR in terms of the value/desirability of outcomes, for, in his formulation of TCR, he just means for his talk of the value/desirability of an outcome to stand for the agent's reasons for wanting that outcome to obtain.

So readers should note that TCR is to be understood in terms of reasons to desire and not necessarily in terms of (impersonal) value or desirability. With that said, I will occasionally revert back to talking about the value/desirability of states of affairs, since this is the language that TCR's critics so often employ. Keep in mind, though, that these critics



mean for value/desirability to somehow include what's "good from a particular individual's point of view," which they equate with "the way she has reason to want things to go" (SCANLON 1998, p. 81). So they are not using words such as 'valuable' and 'desirable' in their ordinary, impersonal senses. The reader should, then, assume that when I revert back to talk of desirability so as to engage TCR's critics on their own terms, I'm using the word 'desirable' as they do, as a kind of short-hand for 'that which the agent has sufficiently weighty reasons to desire'.

§§3.1.2 *Total outcomes as opposed to Intended effects*: Another way that my formulation differs from Scanlon's is that I formulate TCR broadly in terms of reasons to desire total outcomes, and not narrowly in terms of reasons to desire only the results that the agent aims to produce. This, I think, is clearly Scanlon's intent as well, since, as he points out, the "result" of the action must take into account "the action itself" (1998, p. 84). The problem is that his formulation of TCR, as quoted at the beginning of this section, is somewhat unclear on this point in that it refers to "*that* result" [emphasis added], which in this case refers back to the result at which the action was aimed. The problem, then, is that the reasons for performing an action may lie, in part, with the results at which the agent didn't aim. I assume that Scanlon would agree, and, thus, I have formulated TCR in terms of total outcomes as opposed to intended effects.

Note, then, that, as I've formulated TCR, it is not restricted to only the causal consequences of actions.<sup>11</sup> Indeed, it would be odd for the teleologist (which is what I call the proponent of TCR) to exclude in advance from consideration any of the ways that the world might change as a result of an agent's performing an act. For instance, one way the world changes when a subject, S, performs an act,  $a_1$ , is that the world becomes one in which S has performed  $a_1$ . And, as a result of S's having performed  $a_1$ , it may also thereby be true that S has fulfilled her past promise to perform an act of a certain type.<sup>12</sup> Since all these ways in which the world might change could potentially make a difference as to whether or not S has a reason to desire that  $o_1$  obtains and, if so, how much reason, we

<sup>11</sup> Even critics of TCR admit that teleologists can be concerned with more than just the causal consequences of acts. See, for instance, SCANLON 1998 (pp. 80 & 84) and WALLACE 2005 (pp. 26-27).

<sup>12</sup> Thus TCR will not be exclusively forward-looking. The teleologist can hold that S has a reason to desire that  $o_1$  obtains in virtue of the fact that it is true in that world that S has performed  $a_1$  and that her having done so constitutes the fulfillment of her past promise to perform an act of a certain type. For more on this point, see STURGEON 1996, pp. 511-514, and ANDERSON 1996, p. 541.

should formulate TCR, as I have, so that it does not exclude from consideration such possibly relevant non-causal effects.

§§3.1.3 *Reasons (to intend) to act*: Following Scanlon, when I talk about reasons to perform an action, I am, strictly speaking, referring to reasons to intend to perform that action. Most immediately, practical reasoning gives rise not to bodily movements, but to intentions. Of course, when all goes well, these intentions result in some bodily movement, and the end-product is, then, an intentional action. Nevertheless, the most immediate product of practical reasoning is an intention to perform some act, not the act itself (SCANLON 1998, pp. 18-22). Having clarified this, I will, however, sometimes (when it seems not to matter) slip into the more customary way of speaking in terms of reasons for action.<sup>13</sup>

Note, though, that I don't consider facts such as the fact that I will receive some reward or punishment if I intend to  $\varphi$  to constitute genuine reasons for or against intending to  $\varphi$ . The judgment that the intentional content of one's intention to  $\varphi$ , viz.,  $\varphi$ , has certain consequences that one has reason to desire (call this a *content-directed judgment*) is the sort of judgment that can give rise to an intention to  $\varphi$  and that, if true, constitutes a reason to  $\varphi$ . But the judgment that one's having the attitude of intending to  $\varphi$  would have consequences that one has reason to desire (call this an *attitude-directed judgment*) is not the sort of judgment that can give rise to an intention to  $\varphi$ , nor is it, to my mind, the sort of judgment that, if true, constitutes a reason to intend  $\varphi$ .<sup>14</sup> If I'm wrong about this, if the truth of such an attitude-directed judgment does constitute a reason to intend to  $\varphi$ , then I will need to rethink my view that a reason to intend to  $\varphi$  is just a reason to  $\varphi$ , for the fact that I will be rewarded if I intend now to drink some toxin tomorrow is certainly no reason to drink that toxin when tomorrow comes around.<sup>15</sup>

§§3.1.4 *The narrow as opposed to the broad construal of desire*: Although some philosophers (e.g., HEUER 2004, p. 48) take a desire to be nothing more than a disposition to act, where one desires that  $o_i$  obtains if and only if one is disposed to act so as to bring it about that  $o_i$  obtains (that is, to perform  $a_i$ ), I will, following Scanlon, use 'desire' in the more narrow,

<sup>13</sup> As I see it, all talk of reasons for action is really just a somewhat sloppy but more idiomatic way of talking about reasons for intending to act. That is, I don't think that there is any distinction between what we ought, or have reason, to do and what we ought, or have reason, to intend to do. The latter is just a precisification of the former. If I'm wrong about this, then TCR should be taken to be a view about reasons for action.

<sup>14</sup> See SCANLON 2007, esp. pp. 90-91

<sup>15</sup> This example comes from KAVKA 1983.

ordinary sense, such that one desires that  $o_i$  obtains only if one finds the prospect of  $o_i$ 's obtaining in some way attractive or appealing.<sup>16</sup>

On this more narrow interpretation, desiring that  $o_i$  obtains is sufficient for being motivated (to some extent) to perform  $a_i$ , but being motivated to perform  $a_i$  is not sufficient for desiring that  $o_i$  obtains.<sup>17</sup> Thus, in Warren Quinn's famous example of a man who has a compulsive urge that disposes him to turn on every radio he sees despite his failing to see anything appealing about either these acts themselves or their effects (1993, p. 236), we do not have a genuine case of desire—at least, not in the sense that I will be using the term 'desire'. As I see it, then, having a desire involves a complicated set of dispositions to think, feel, and react in various ways (SCANLON 1998, p. 21). A person who desires that  $o_i$  obtains will find the prospect of its obtaining appealing, will be to some extent motivated to perform  $a_i$ , and will, perhaps, have her attention directed insistently toward considerations that present themselves as counting in favor of  $o_i$ 's obtaining (SCANLON 1998, p. 39).<sup>18</sup>

Unlike Scanlon, though, I don't think that "desiring something involves having a tendency to see something good or desirable about it" (1998, 32). This suggests that preferring  $o_i$  to  $o_j$  involves having a tendency to see  $o_i$  as better than (or preferable to)  $o_j$ . But I don't think that can be right, as I can, even other things being equal, prefer the outcome in which I am saved to the outcome in which five others are saved without having any tendency to see the former as better than the latter. Of course, I might rightly think that the former is better *for* me, but that's not the same as thinking that it would be better if I was the one whom was saved. I think, then, it's more accurate to say that desiring something typically involves having a tendency to see something about it as providing one with a reason to desire it—likewise, for preferring.

§§3.1.5 *TCR as opposed to just TCR-1+2*: To be a teleologist, it is not enough to accept the bi-conditional that is entailed by the conjunction of TCR-1 and TCR-2 (call this bi-conditional "TCR-1+2"); the teleologist must

<sup>16</sup> To find  $o_i$ 's obtaining in *some way* attractive or appealing, one needn't have all that  $o_i$ 's obtaining entails conscious before the mind. Ana might know that the only way to ensure that her daughter excels in school is to hire Bill (a tutor) and, consequently, find the outcome resulting from her hiring Bill in this respect appealing.

<sup>17</sup> Because one can be motivated (to some extent) to perform  $a_i$  without being sufficiently motivated to perform  $a_i$ , desiring that  $o_i$  obtains does not necessarily result in an intention to perform  $a_i$ . After all, one can have conflicting motives.

<sup>18</sup> I am not sure whether the last of these three is essential to desiring, as the qualifier 'perhaps' is meant to indicate. For reasons to doubt that it is essential to desiring in the ordinary sense, see CHANG 2004, esp. pp. 65-66.

accept TCR-3 as well. Of course, TCR-3 is but one of three possible explanations for the truth of the bi-conditional stated by TCR-1+2. To illustrate, let 'PER' stand for 'S has more reason to *perform*  $a_i$  than to perform  $a_j$ ' and let 'DES' stand for 'S has more reason to *desire* that  $o_i$  obtains than to desire that  $o_j$  obtains'. The three possible explanations for TCR-1+2—that is, for 'PER if and only if DES'—are: (i) 'PER, because DES', (ii) 'DES, because PER', or (iii) both 'PER, because BET' and 'DES, because BET'—where, for instance, 'BET' might stand for ' $o_i$  is *better* than  $o_j$ '.<sup>19</sup> In defending TCR, I must not only defend TCR-1+2, but also argue that it is explanation *i* as opposed to either explanation *ii* or explanation *iii* that explains TCR-1+2; that is, I must defend TCR-3 in addition to both TCR-1 and TCR-2. But before I proceed to defend TCR, I will first try to clear up some actual and potential misconceptions about the view.

### §3.2 Clearing up some misconceptions about TCR

There are a number of misconceptions about TCR that have led philosophers to reject TCR for mistaken reasons. Below, I try to clear up some of these misconceptions.

§§3.2.1 *TCR is compatible with value concretism*: Although Scanlon (1998, 79-81) lumps the two together, TCR is distinct from, and independent of, *value abstractism*: the view that the sole or primary bearers of intrinsic value are certain abstracta—facts, outcomes, states of affairs, or possible worlds.<sup>20</sup> On value abstractism, there is only one kind of value, the kind that is to be promoted, and so the only proper response to value is to desire and promote it, ensuring that there is as much of it as possible. The contrary view—the view that the fundamental bearers of intrinsic value are concrete entities (e.g., persons, animals, and things)—is called *value concretism*.<sup>21</sup>

Contrary to what Scanlon and others (e.g., ANDERSON 1993; WALLACE 1995) have claimed, there is no reason why the teleologist cannot

<sup>19</sup> These three explanations are analogues of the three possible causal explanations for a correlation between events *a* and *b*: (1) *a* causes *b*, (2) *b* causes *a*, or (3) *a* and *b* have a common cause. I thank Mark Schroeder and Shyam Nair for pointing out the need to consider such common "cause" explanations. And I thank Schroeder for suggesting that someone might take ' $o_i$  is *better* than  $o_j$ ' to be the common "cause."

<sup>20</sup> This is true whether TCR is to be formulated in terms of how much reason the agent has to desire the available outcomes or in terms of how valuable/desirable (in the ordinary, agent-neutral sense) the available outcomes are.

<sup>21</sup> I borrow the terms *concretism* and *abstractism* from TÄNNISJÖ 1999.

accept value concretism.<sup>22</sup> Indeed, the teleologist can accept all of the following: (a) that concrete entities—persons, animals, and things—are the primary bearers of intrinsic value; (b) that states of affairs generally have only extrinsic value in that they generally have no value apart from our valuing concrete entities;<sup>23</sup> (c) that “our basic evaluative attitudes—love, respect, consideration, honor, and so forth—are non-propositional...attitudes we take up immediately toward persons, animals, and things, not toward facts” (ANDERSON 1993, p. 20); and (d) that value and our valuations are deeply pluralistic, that there are many ways that we experience things as valuable (e.g., as interesting, admirable, beautiful, etc.) and that there are many different kinds of value as well as different modes of valuing that are appropriate to each (e.g., “beautiful things are worthy of appreciation, rational beings of respect, sentient beings of consideration, virtuous ones of admiration, convenient things of use”—ANDERSON 1993, p. 11). TCR is compatible with all of (a)-(d), as I will now explain.

As rational *beings*, we appropriately respond to different sorts of things with different sorts of attitudes. We appropriately respond to beautiful objects by appreciating them, we appropriately respond to rational persons by respecting them, and we appropriately respond to desirable states of affairs (desirable in the ordinary, agent-neutral sense) by desiring their actualizations—at least, that’s how we appropriately respond to them when we don’t have weightier agent-relative or time-relative reasons to desire that they not be actualized. As rational *agents*, though, it is only the last of these three that is pertinent, for, as agents, we can effect only outcomes. We cannot effect valuable entities. A concrete entity is not the sort of thing that we can bring about or actualize through our actions.

<sup>22</sup> Elizabeth Anderson is, as I see it, another leading critic of TCR. Although she uses the term ‘consequentialism’ as opposed to ‘teleology’, she defines ‘consequentialism’ so broadly (see 1993, 30-31) that it is, in spirit, equivalent to TCR. She says, for instance, “consequentialism specifies our rational aims, and then tells us to adopt whatever intentions will best bring about those aims” (ANDERSON 1996, p. 539), which is exactly what TCR tells us to do. Thus, as Anderson uses the term, ‘consequentialism’ refers not to a moral theory but to a conception of practical reasons that is roughly equivalent to TCR. In certain passages, Anderson, like Scanlon, defines ‘consequentialism’ in terms of value as opposed to reasons to desire. But, as with Scanlon, this is only because she talks as if intrinsic goods include both what’s good for an individual and what’s good relative to an individual (ANDERSON 1993, pp. 30-31).

<sup>23</sup> The reason for the qualifier ‘generally’ in item *b* is that Anderson does allow for the possibility that a state of affairs can have intrinsic value if it is one that is intrinsically interesting. Anderson says, “Interest does seem to be an evaluative attitude that can take a state of affairs as its immediate and independent object. This is an exception to the general rule that states of affairs have no intrinsic value” (1993, p. 27).

Of course, we can act so as to bring it about that, say, a rational person exists or that our actions express our respect for rational persons, but these are states of affairs, not concrete entities.

As agents, then, we have the ability to actualize only certain possible worlds or states of affairs. Indeed, purposive action must aim at the realization of some state of affairs. So the teleologist can admit that we have reasons to have all sorts of different attitudes, including reasons to have certain non-propositional attitudes (such as, respect) toward various concrete entities (such as, rational persons). But the teleologist will insist that when it comes to the particular attitude of intending to act in some way, the reasons for having this attitude must always be grounded in the reasons that the agent has to desire that certain possible worlds or states of affairs be actualized. It is a mistake, however, to think that the teleologist is, in addition, committed to the denial of any of claims (a)-(d) above.<sup>24</sup>

§§3.2.2 *TCR is compatible with appropriately valuing goods such as friendship*: Another common misconception regarding TCR is that it is incompatible with the thought that, with respect to goods such as science and friendship, taking them to be valuable is not simply, or even primarily, a matter of promoting certain states of affairs (cf. SCANLON 1998, p. 88). Take friendship. TCR does not imply, for instance, that the only reasons provided by my friend and our friendship are reasons to promote certain states of affairs. The teleologist can accept that I have reasons to care about my friend, to empathize with her pain, to take joy in her successes, etc. and that these are not reasons to promote certain states of affairs, but rather reasons to have certain non-propositional attitudes and feelings. TCR is, then, compatible with the thought that what lovers, friends, and family members value, fundamentally, is *each other* as opposed to certain states of affairs.

TCR is also compatible with the thought that a person who values friendship will see that what she has reason to do, first and foremost, is to be a good friend to her current friends and that these reasons are weightier than whatever reasons she has to cultivate new friendships or to foster good friendship relations among others (cf. SCANLON 1998, pp. 88-89). The teleologist can even hold that my friendships generate agent-centered restrictions on my actions (cf. ANDERSON 1993, pp. 73-74), such that I have

<sup>24</sup> This point is not particularly new, although it bears repeating given the stubborn persistence of this misconception. Others who have made essentially the same point include ARNESON 2002 and STURGEON 1996.

more reason to refrain from betraying one of my own friends than to prevent more numerous others from betraying theirs.

This is all possible given that TCR allows for agent-relative reasons. If there were only agent-neutral reasons, (e.g., agent-neutral reasons to promote friendships and to prevent friends from betraying one another), then I would often have sufficient reason to neglect one of my current friendships if I could thereby cultivate two or more new ones, and I would often have sufficient reason to betray one of my own friends if I could thereby prevent more numerous others from betraying theirs. But, as even critics of TCR admit (e.g., SCANLON 1998, p. 81), TCR is compatible with the existence of agent-relative reasons and thus with the idea that whereas *you* will have more reason to prefer that *your* friends are not betrayed, *I* will have more reason to prefer that *my* friends are not betrayed.<sup>25</sup> And such agent-relative reasons to prefer the possible world in which your friends as opposed to my friends are betrayed will, given TCR, generate agent-relative reasons for me to refrain from betraying one of my own friends even for the sake of preventing you from betraying two of yours.<sup>26</sup> And TCR certainly allows for the possibility that such reasons will be decisive and thereby generate an agent-centered restriction against betraying one's own friends even for the sake of preventing more numerous others from betraying theirs.

What's more, TCR is compatible with the claim that I should not abandon my current friends even for the sake of cultivating more numerous new friendships, for the teleologist can hold that I currently have good time-relative reasons for preferring the preservation of my current friendships to the creation of otherwise similar new friendships given the shared history that I have with my current friends and the lack of any shared history (at present) with those possible future friends (HURKA 2006, p. 238).<sup>27</sup> And not only can the teleologist accept that one should not destroy one's own current friendships for the sake of creating more

<sup>25</sup> TCR would be incompatible with the existence of agent-relative reasons only if TCR were to be formulated in terms of how valuable/desirable (in the ordinary, agent-neutral sense) the available outcomes are. But, as I've shown, neither Scanlon nor Anderson thinks that this is the way to formulate TCR.

<sup>26</sup> Similarly, the teleologist can even hold that I have more reason to prefer that my friends are betrayed by *you* than to prefer that they are betrayed by *me*. And such agent-relative reasons will, given TCR, generate agent-relative reasons for me to refrain from betraying one of my friends even for the sake of preventing you from betraying two of my friends.

<sup>27</sup> As mentioned in note 10, TCR should, then, be formulated as follows: S has more reason to perform  $a_i$  at  $t_i$  than to perform  $a_j$  at  $t_i$  if and only if, and because, S has more reason, at  $t_i$ , to desire that  $o_i$  obtains than to desire that  $o_j$  obtains.

numerous future friendships for oneself or for others, but also that, out of respect for friendships generally, one should not destroy someone else's friendship for the sake of preventing numerous others from doing the same. Again, because the teleologist can hold that there are agent-relative reasons for preferring one possible world to another, the teleologist can hold that I should prefer the state of affairs in which, say, five others each destroy someone else's friendship to the state of affairs in which I myself destroy someone else's friendship.

§§3.2.3 *TCR is compatible with the view that attitudes such as belief and blame are rationally justified on non-pragmatic grounds*: Elizabeth Anderson has claimed that the teleologist is committed to the implausible view that all attitudes (blaming, believing, intending, etc.) are rationally justified on pragmatic grounds—that is, on the grounds that the agent's having the given attitude would have desirable consequences. For instance, she claims that the teleologist (or what she calls the "consequentialist") must hold that beliefs "are justified to the degree that they bring about better states of affairs into existence" (ANDERSON 1993, p. 39). This is mistaken for at least three reasons.

First, note that the teleologist does not even hold that whether one is rationally justified in having the intention to perform  $a_i$  is a function of the desirability of the consequences of one's having this attitude. TCR does not imply, for instance, that the fact that an evil demon has threatened to produce undesirable consequences unless you intend to perform  $a_i$  gives you a reason to intend to perform  $a_i$ . TCR implies that it is the fact that your *performing*  $a_i$  will have desirable consequences (or, as I would prefer to say, consequences that you have sufficiently weighty reasons to desire), not the fact that your *intending to perform*  $a_i$  will have desirable consequences, that provides you with a reason to intend to perform  $a_i$ .<sup>28</sup> Thus, it is the consequences of *the act*, not of *the intention*, that is relevant on TCR. And so there is no way to generalize from the claim that intending to do  $a_i$  is justified on the grounds that *performing*  $a_i$  is instrumental in bringing about certain desirable consequences to the claim that attitudes, such as believing that  $p$ , are justified on the grounds that *having this attitude*

<sup>28</sup> To see that these can come apart, consider that one can, say, knock over a bucket without forming the intention to do so and that one can form the intention to do so without succeeding. On TCR, the reasons that one has to intend to knock over a bucket depend on what the world would be like if one knocked over the bucket (whether or not one does so intentionally), not on what the world would be like if one formed the intention to do so (whether or not one succeeds).



is instrumental in bringing about certain desirable consequences. It is just not analogous.

Second, even if it were true that the teleologist thought that the rationality of intending to perform  $a_i$  depends on the desirability of the consequences of having this attitude, it is not clear why the teleologist would be committed to the more general claim that this is true of *all* attitudes. Why can't the teleologist hold that what is true of intending to act is not true of other attitudes?<sup>29</sup>

Third, it seems that Anderson must be mistakenly assuming that mental attitudes such as blaming and believing are actions in the relevant sense, as this seems to be the only possible explanation for why she thinks that the teleologist is committed to the view that blaming and believing are rationally justified on the grounds that performing these "actions" would have desirable consequences. But, unlike intentional actions, we don't "perform" the "act" of blaming and believing at will—that is, by intending to blame or by intending to believe. Take belief, for instance. We often form beliefs involuntarily in response to our perceptual experiences. These belief formations are not intentional actions, for we do not form these beliefs as the result of our intending to form them. Since blaming and believing are not intentional actions, they are not the sorts of things to which TCR applies.

So the teleologist can accept, contrary to what Anderson claims, that whether a person has sufficient reason to believe that  $p$  depends on what her evidence is for the truth of  $p$  and not on the desirability of the consequences of her believing that  $p$ . What's more, the teleologist can accept that whether blaming someone is rationally justified depends only on whether there is sufficient reason to blame that someone. Of course, the teleologist is committed to the view that whether one is rationally justified in intending to act so as to criticize, punish, or otherwise censure someone does depend on the consequences of such an act. But this is distinct from the attitude of blaming someone for her actions, i.e., that of feeling guilt, resentment, or indignation in response to her actions (LEVY 2005, p. 2). One can blame a person for her wrongdoing without intending to act so as to criticize, punish, or otherwise censure her.

<sup>29</sup> Or if one prefers to talk about reasons for action as opposed to reasons for intending to act, why can't the teleologist hold that although reasons for action are a function of our reasons for desiring its consequences, neither reasons for believing nor reasons for blaming are a function of our reasons for desiring the consequences that stem from our having these attitudes?

§§3.2.4 *TCR is compatible with passing the normative buck from value to reasons*: Lastly, it should be noted that TCR is compatible with a buck-passing account of value, where the domain of reasons is taken to be explanatorily prior to the domain of value (cf. WALLACE 2005)—thus the normative buck is passed from value to reasons. There is potential for confusion here, for Scanlon sometimes formulates the buck-passing account of value specifically in terms of reasons for *action*, and TCR does treat evaluative reasons (e.g., reasons to desire) as explanatorily prior to practical reasons (i.e., reasons for action).<sup>30</sup> Nevertheless, TCR is compatible with the following more general formulation of the buck-passing account of value: *x*'s being good or valuable is just the purely formal, higher-order property of having other properties that provide sufficiently weighty reasons of the right kind to respond favorably toward *x*. This leaves open what the relevant favorable response will be, and, arguably, it will vary depending on the kind of thing that's being evaluated. In the case of a rational person, the relevant response might be to have respect for that person. In the case of a valuable state of affairs, the relevant response might be to desire that it obtains. If this is right, then TCR is perfectly compatible with the buck-passing account of value. An agent-neutral teleologist could, for instance, hold that reasons for actions are a function of the value (or the desirability) of the states of affairs that those actions produce, but still ultimately pass the normative buck back to reasons by claiming that what it is for a state of affairs to be valuable/desirable (in the ordinary, agent-neutral sense) is for it to have the purely formal, higher-order property of having other properties that provide sufficiently weighty reasons (of the right kind) to *desire* it.<sup>31</sup>

### §3.3 Scanlon's putative counterexamples to TCR

Besides these misconceptions, some philosophers are led to reject TCR given putative counterexamples. Scanlon, for instance, uses examples to argue that "many reasons bearing on an action concern not the desirability of outcomes but rather the eligibility or ineligibility of various other reasons..., and judging that a certain consideration does not count as a

<sup>30</sup> Scanlon says, for instance, that "to call something valuable is to say that it has other properties that provide reasons for behaving in certain ways with regard to it" (1998, p. 96).

<sup>31</sup> The agent-neutral teleologist holds that the best outcome (in the ordinary, agent-neutral sense) available to a given agent is necessarily the one she ought to prefer to all other available alternatives, either because fittingness reasons for preferring one outcome to another are the only reasons for preferring one outcome to another or because fittingness reasons always override non-fittingness reasons.

reason for action is not equivalent to assigning negative intrinsic value to the occurrence of actions based on this reason" (1998, p. 84). Unfortunately, it is not clear what he is trying to establish by this. One possibility is that he is pointing out that there are sometimes reasons for *believing* (he says "judging") that a certain fact does not count as a reason to perform a given act and that these *epistemic* reasons (i.e., reasons for belief) do not concern the desirability of outcomes. But, then, this is no counterexample to TCR. To provide a counterexample to TCR, Scanlon must provide an example in which we are compelled to think that there are *practical* reasons (reasons for action) that do not concern the desirability of outcomes, not an example in which we are compelled to think that there are *epistemic* reasons (reasons for belief) that do not concern the desirability of outcomes. As we saw in §3.2.3, the teleologist can accept that reasons for belief do not concern the desirability of outcomes. This, then, is clearly not the most charitable way to interpret Scanlon.

So let us consider his actual examples. In Chapter 1, Section 10, he provides examples where an agent "judge[s] one consideration, C, to be a reason for taking another consideration, D, not to be relevant to... [his or her] decision to pursue a certain line of action" (1998, p. 51). What he gives us, then, are not reasons that concern the eligibility or ineligibility of various other reasons, but rather reasons that concern whether or not one ought to *take* various other reasons into account in one's deliberations, and taking other reasons into account in one's deliberation is itself an action.<sup>32</sup> Moreover, the reasons for or against taking other reasons into account in one's deliberations do seem to concern the comparative desirability of the outcomes associated with doing so versus not doing so. Indeed, in the sorts of examples that Scanlon provides, the reason not to take a certain consideration into account in one's deliberations is that doing so would have an undesirable effect (or, at least, an effect that one has good reason to want to avoid). And so, Scanlon's putative "counterexamples" are not counterexamples after all.

<sup>32</sup> The verb 'take' has many senses. In one sense, to take something to be irrelevant is to regard or treat it as irrelevant with the implication that this may be contrary to fact. This is how I am interpreting Scanlon here. Thus I am assuming that, by the phrase "taking another consideration, D, not to be relevant," he means "treating another consideration, D, as being irrelevant (whether or not it is, or is believed to be, relevant)." If instead Scanlon is using his phrase to mean "believing D not to be relevant," then he is not talking about reasons for actions (or, as he says, "reasons bearing on an action"), but is instead talking about reasons for belief. If he is talking about such epistemic reasons, then, as I explained above, this poses no problem for TCR, which instead concerns practical reasons.

It seems, then, that Scanlon may be conflating reasons for and against performing an act,  $a_i$ , with reasons for and against taking into account various considerations when deciding whether to perform  $a_i$ . But the act of taking into account (or ignoring) various considerations when deciding whether to perform  $a_i$  is not the act of performing  $a_1$ , but instead the act of performing some other act, say,  $a_2$ . So let  $a_2$  be the act of taking into account certain considerations when deciding whether to perform  $a_1$ .<sup>33</sup> It seems that when we consider whether or not to perform  $a_2$ , we should, as TCR implies, consider the agent's reasons for and against desiring that  $o_2$  obtains. Thus TCR does not deny that the reasons for and against performing  $a_2$  (i.e., taking into account certain considerations when deciding whether to perform  $a_1$ ) may have nothing to do with the reasons for and against desiring that  $o_1$  obtains. What it denies is only that the reasons for and against performing  $a_2$  may have nothing to do with the reasons for and against desiring that  $o_2$  obtains. And none of Scanlon's examples repudiate this, as I'll show presently.

Consider one of Scanlon's main examples, one where I have met someone for a game of tennis. Assume that I have determined that there are no strong reasons for or against my playing to win so that whether I have reason to play to win just depends on what I would enjoy doing at the moment. And assume that what I would enjoy most at this moment is playing to win and so this is what I have decided to do. Given all this, Scanlon claims, contrary to TCR, that the fact that my succeeding in making a certain strategic shot might make my opponent feel crushed or disappointed just is not relevant to whether or not I should make the shot.

I don't think that we should accept this claim. Now, Scanlon never says what consideration is the reason for me to ignore this other consideration (i.e., that my opponent might feel crushed) when deciding whether or not to make this strategic shot, but clearly it is some pragmatic

<sup>33</sup> I'm speaking a bit loosely when I treat taking into account (or ignoring) some consideration as itself a voluntary act, for, admittedly, whether one takes into account (or ignores) some consideration is very rarely, if ever, directly under one's volitional control. So, one doesn't so much intend to take into account (or ignore) some consideration as intend to do that which will cause oneself to take into account (or ignore) that consideration. For instance, if I want to take into account certain considerations, I will intentionally focus my attention on considerations of that sort. And if, instead, I want to ignore those considerations, I will intentionally focus my attention elsewhere, purposely and immediately diverting my attention to something else whenever a consideration of that sort creeps into my consciousness. So the expression 'taking into account (or ignoring) certain considerations' should be taken as a convenient shorthand for 'doing that which will cause oneself to take into account (or ignore) certain considerations'.

consideration such as the fact that it is very difficult, if not impossible, to play to win while taking these sorts of considerations into account. Or maybe it's the fact that one cannot enjoy a good competitive game if one is constantly worrying about whether one's opponent's feelings might get hurt. But notice that we are now appealing to the desirability of taking my opponent's feelings into account, which is precisely what the teleologist holds to be relevant. The fact that we are not appealing to the desirability of my making the strategic shot and weighing the desirability of the outcome in which I make the shot against the desirability of the outcome in which I don't is neither here nor there. So it seems that this example shows only that the desirability or undesirability of embarrassing my opponent is not relevant to whether or not I should ignore such considerations. It does not, however, show that the desirability or undesirability of embarrassing my opponent is not relevant to whether or not I should make the shot.

Perhaps, even this is not the most charitable way of interpreting Scanlon. Perhaps, we should interpret Scanlon to be offering this example, not as a counterexample, but rather as the first step in an argument against TCR. That is, Scanlon could be arguing as follows:

- 3.1 In many cases, we do not, when deliberating about what to do, treat the fact that an act would have certain desirable/undesirable consequences as a reason for/against performing it. For instance, when playing to win, we do not treat the fact that taking a certain shot would embarrass one's opponent (an admittedly undesirable consequence) as a reason against taking the shot.
- 3.2 Our deliberative experience in such cases fits better with (*a*) the contention that these are cases in which the agent treats as non-reasons facts that are actually non-reasons than it does with (*b*) the contention that these are cases in which, on pragmatic grounds, the agent acts so as to cause herself to treat as non-reasons facts that are actually reasons (reasons with actual weight).<sup>34</sup>
- 3.3 Therefore, absent some other reason for accepting *b*, we should accept *a*, which supports  $\sim$ TCR.<sup>35</sup>

<sup>34</sup> In calling these considerations 'non-reasons', I mean to allow that these considerations may count as genuine reasons in other contexts. The assumption, here, is only that due to their being silenced, undermined, or bracketed off, these considerations have ceased to be reasons in the context at hand.

<sup>35</sup> Something along these lines was suggested to me by Peter de Marneffe.

Unfortunately, though, Scanlon never even considers *b* and so never explains why he thinks that the phenomenology of our actual deliberations supports *a* as opposed to *b*. Perhaps, he just thinks that this obvious. If so, then all I can do is cite my dissent. When I reflect on my first-person deliberations in such cases, it feels to me like I'm deciding on pragmatic grounds to ignore certain relevant reasons. To borrow another one of Scanlon examples: when I serve on a search committee and ignore how I would personally benefit from our hiring the candidate who specializes in moral philosophy, this feels like a case where I am choosing to ignore a genuine reason that I have to recommend the candidate who specializes in moral philosophy to the department. I choose to ignore the self-interested reason that I have for doing so, because I recognize that I have better reason to want to live up to the standards associated with the role of being a good search committee member, and those standards require me to make my recommendation on the basis of what would be best for our program as opposed to what would be best for me.

So when I consider the phenomenology in these cases, they seem to me to support *b*, and Scanlon gives me no reason to think that I'm mistaken in this, as he fails to even consider *b*, let alone explain why he thinks that the phenomenology supports *a* instead. And, even if I'm wrong about the phenomenology, Scanlon's argument fails to establish that TCR is false, for, as we'll see presently, there are reasons to accept TCR and thus *b* even if the phenomenology supports *a* instead. After all, things are not always as they seem. Thus, even if I'm wrong and the phenomenology fits better with *a*, we should accept *a* only if there are no better reasons for accepting *b* instead.

### §3.4 Arguments for TCR

Having argued that many of the reasons that philosophers have given for rejecting TCR are not in fact good reasons, I will now argue that there are good reasons to accept each of TCR's three claims. Furthermore, I will argue that there are good reasons for accepting TCR that go beyond the reasons for accepting each of its three claims.

§§3.4.1 *In defense of TCR-1*: According to TCR-1, if *S* has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains, then *S* has more reason to perform  $a_i$  than to perform  $a_j$ . To deny TCR-1 is to hold that the following is possible: an agent ought to prefer the outcome in which she has performed  $a_i$  (viz.,  $o_i$ ) even though  $a_i$  is not the alternative that she

ought to perform. Given such a possibility, the agent ought to hope that she performs the alternative (viz.,  $a_i$ ) that she ought not to perform. And if, in spite of hoping that she'll perform  $a_i$ , she performs  $a_j$ , she ought to wish that she had performed  $a_i$  instead. I find this counterintuitive. Surely, agents shouldn't hope that they'll act as they shouldn't. Nor should they wish that they had acted as they shouldn't have acted. To avoid such counterintuitive implications, we should accept TCR-1.

To illustrate how the denial of TCR-1 can lead to such counterintuitive implications, imagine that I'm picnicking on a hill beside a lake and that I must decide whether or not to run as fast as I can down the hill to the lake's only boat dock. Suppose that Richard and his two daughters are complete strangers to me, that my daughter and Richard's two daughters are all drowning, that there's only enough time to rescue either my daughter or Richard's two daughters (as they're at separate ends of the lake), that there's only one boat tied to the boat dock, and that there's no other way to rescue the children. Further suppose that if I run as fast as I can to the boat dock, I'll get there first and rescue my daughter. If, however, I don't run as fast as I can, Richard will get there first and rescue his two daughters. Assume that everything else is equal.

Let's stipulate that  $o_1$ —the possible world in which it's true to say that I ran as fast as I could and rescued my daughter—will obtain if and only if I perform  $a_1$ —the act of running as fast as I can down to the boat dock. And let's stipulate that  $o_2$ —the possible world in which it's true to say that I refrained from running as fast as I could and thereby allowed Richard to rescue his two daughters—will obtain if and only if I perform  $a_2$ —the “act” of refraining from running as fast as I can down to the boat dock.<sup>36</sup> Lastly, let's assume, for the sake of argument, that  $o_2$  is, impersonally speaking, better than  $o_1$ , but that, in this instance, I ought not to do what would bring about the impersonally best outcome (viz.  $a_2$ ). That is, let's assume that I should perform  $a_1$ , thereby rescuing my daughter.

Interestingly, if we deny TCR-1, the following is possible: I ought, all things considered (that is, have most reason, all things considered), to prefer  $o_2$  to  $o_1$  even though I ought, as stipulated, to perform  $a_1$  as opposed to  $a_2$ . Given such a possibility, I ought to hope that I perform the act that I

<sup>36</sup> For the sake of simplifying the discussion, I'm assuming that there are only two act-tokens available to me: one that falls under the description of my running as fast as I can down to the boat dock and another that falls under the description of my refraining from running as fast as I can down to the boat dock. If it's necessary, the reader can assume that these are the only two alternative act-tokens available to me, because God will immediately cause me to cease to exist should I attempt to perform any other act-token.

ought not to perform. After all, it seems that if I ought, all things considered, to prefer the outcome in which I have performed  $a_2$  (viz.,  $o_2$ ), an outcome that will obtain only if I perform  $a_2$ , then I ought to hope that I will perform  $a_2$ . Yet  $a_2$  is the alternative that I ought not to perform. Thus, on this possibility, I ought to hope that I will act as I ought not to act. Again, I find this counterintuitive. It seems much more plausible to accept TCR-1 and thereby deny such possibilities. We can deny such possibilities by accepting TCR-1 and claiming that since I ought to perform  $a_1$  as opposed to  $a_2$ , it can't be that I ought to prefer  $o_2$  to  $o_1$ .

Of course, I've stipulated that  $o_2$  is, impersonally speaking, better than  $o_1$ . And, given this, an impartial spectator ought to prefer  $o_2$  to  $o_1$ . But just because an impartial spectator ought to prefer  $o_2$  to  $o_1$ , we shouldn't infer that I ought to prefer  $o_2$  to  $o_1$ . After all, I differ from an impartial spectator in that I'm neither impartial nor a mere spectator. Unlike an impartial spectator, I bear certain agential relations to these two outcomes (i.e., I'm the agent whose actions determine which outcome will obtain), and I bear a special relation to my daughter, who will be better off if  $o_1$  obtains. Given these relations and, in particular, the special relationship that I have with my daughter, I ought to prefer  $o_1$  to  $o_2$  despite the fact that  $o_2$  is, impersonally speaking, better than  $o_1$ .

So, although I have good agent-neutral reasons for preferring  $o_2$  to  $o_1$ , it seems that I have even better agent-relative reasons for preferring that my daughter is saved. To deny this is to allow that I can have decisive reasons to act so as to save my own daughter even though I have decisive reasons to prefer the outcome in which I have instead acted so as to allow Richard to save his two daughters. Such a claim is, I've suggested, counterintuitive. And if we want to avoid such counterintuitive implications, then we should hold that an agent's reasons for performing an act must always track her reasons for preferring the outcome in which she has performed that act to the outcome in which she has refrained from performing that act. And, thus, we should think that it will never be the case that an agent has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains but more reason to perform  $a_j$  than to perform  $a_i$ . That is, we should accept TCR-1.

§§3.4.2 *In defense of TCR-2*: Similar considerations count in favor of TCR-2. According to TCR-2, if S has more reason to perform  $a_i$  than to perform  $a_j$ , then S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains. To deny TCR-2 is to hold that the following is possible: an agent ought to perform  $a_i$  even though the outcome in which she has performed  $a_i$  (viz.,  $o_i$ ) is not the outcome that she ought to prefer. Given



such a possibility, the agent ought to perform  $a_i$  even though she ought not to hope that she will perform  $a_i$ . And if, as it turns out, she fails to perform the act that she ought to perform (viz.,  $a_i$ ), she ought not to wish that she had. I find this counterintuitive as well. Surely, agents should hope that they will act as they ought to act and should wish that they had acted as they should have acted. To avoid such counterintuitive implications, we should accept TCR-2.

To illustrate how the denial of TCR-2 can lead to such counterintuitive implications, imagine that I must decide whether or not to lie to Henri about Ida's whereabouts. Assume that Henri is looking to kill Ida, that Ida is completely innocent, that Henri will fail to kill Ida only if I lie to Henri, and that I know all of this. Let's stipulate that  $o_3$ —the possible world in which it's true to say that I have lied to Henri and thereby prevented him from killing Ida—will obtain if and only if I perform  $a_3$ —the act of lying to Henri. And let's stipulate that  $o_4$ —the possible world in which it's true to say that I have refrained from lying to Henri and thereby allowed him to kill Ida—will obtain if and only if I perform  $a_4$ —the "act" of refraining from lying to Henri.<sup>37</sup> Assume that everything else is equal. Lastly, assume, for the sake of argument, that I ought to prefer  $o_3$  to  $o_4$ , for the agent-neutral reasons that I have for preferring the outcome that's least bad (in this case, the one in which I have lied to Henri, thereby preventing him from killing Ida) decisively oppose the agent-relative reasons that I have for preferring the outcome in which I have refrained from treating anyone as a mere means (in this case, the one in which I have refrained from lying to Henri as a means to preventing him from killing Ida).

Interestingly, the denier of TCR-2 can say that although I ought to prefer the outcome in which I have lied to Henri, I ought, as Kant claimed, to refrain from lying to Henri. But if we really believe that I ought to prefer  $o_3$  to  $o_4$ , why would we insist that I ought to perform  $a_4$  as opposed to  $a_3$ ? If we insist that such is the case, then we must hold that I ought to perform the alternative that I ought to hope that I do not perform. It seems much more plausible to accept TCR-2 and accept that if I ought to prefer the outcome in which I have lied to Henri, then I ought, contrary to Kant, to lie to Henri.

So to preserve the intuitions both that an agent ought to hope that she will perform an act if and only if it's the act that she ought to perform

<sup>37</sup> Here too, I'm assuming, merely for sake of simplifying the discussion, that there are only two act-tokens available to me: one that falls under the description of my lying to Henri and another that falls under the description of my refraining from lying to Henri.

and that an agent ought to wish that she had acted differently if and only if she didn't act as she should have, we should accept both TCR-1 and TCR-2. That is, we should accept TCR-1+2: S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains if and only if S has more reason to perform  $a_i$  than to perform  $a_j$ .

§§3.4.3 *In defense of TCR-3:* As I noted in §§3.1.5, the teleologist is committed not only to TCR-1+2, but also to the right-hand side of that bi-conditional having explanatory priority. Thus, the teleologist must defend TCR-3: if S has more reason to perform  $a_i$  than to perform  $a_j$ , then this is in virtue of the fact that S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains.

In defense of TCR-3, I will argue that it is more plausible than its alternatives. If we let 'PER' stand for 'S has more reason to *perform*  $a_i$  than to perform  $a_j$ ' and let 'DES' stand for 'S has more reason to *desire* that  $o_i$  obtains than to desire that  $o_j$  obtains', then TCR-3 is the view that, in every instance, 'PER, because DES'. And the three possible alternatives to this view are: (1) in every instance, 'DES, because PER'; (2) in some but not all instances, 'PER, because DES', and, in every other instance, 'DES, because PER'; and (3) in every instance, both 'PER, because BET' and 'DES, because BET' — where, for instance, 'BET' might stand for ' $o_i$  is *better* than  $o_j$ '.

So the first alternative to TCR-3 is to hold that what explains the truth of the bi-conditional 'PER if and only if DES' is that, in every instance, 'DES, because PER', not 'PER, because DES', as TCR-3 supposes. This is to hold that when S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains, this is *always* because S has more reason to perform  $a_i$  than to perform  $a_j$ . But this is clearly false, as one simple example is sufficed to show. Suppose that  $a_5$  is the act of putting my money into a savings account that yields 5% annually and that  $a_6$  is the act of putting my money into a savings account that yields 3% annually. Assume that both savings accounts are otherwise equal and that, other things being equal, I am better off putting my money in a higher-yielding savings account. In this case, it is clear that what explains the fact that I have more reason to perform  $a_5$  than to perform  $a_6$  is the fact that I have more reason to desire that  $o_5$  obtains than to desire that  $o_6$  obtains. The same would seem to hold for any other purely prudential choice, for surely no one would argue that the reason I have for preferring the outcome in which I am prudentially better off is that this is the outcome in which I did what was prudent (i.e., what I had most prudential reason to do). We should, then, reject the first alternative.

Perhaps, this is too quick. After all, Elizabeth Anderson has argued that an action is rational only if it adequately expresses one's rational evaluative attitudes (e.g., respect, consideration, and appreciation) toward persons, animals, and things. And, on her view, it is rational to adopt the aim (or the end) of bringing about some state of affairs only if the act that would bring it about is itself rational. For instance, on her view, I have reason to want, and to aim to bring about, the state of affairs in which I spend time alone with wife only because it's rational for me to act in ways that express my love for my wife. On Anderson's view, then, whether I have reason to want, and to adopt as my end, the state of affairs in which I spend time alone with my wife depends on whether it is rational for me to act so as to adequately express my love for her. And, perhaps, Anderson would similarly argue that whether I have reason to want, and to adopt as my end, the state of affairs in which I'm better off depends on whether it is rational for me to act so as to make myself better off.

Anderson's view, thus, seems contrary to TCR-3 in that it holds that the rationality of desiring and pursuing some end,  $e_1$ , can depend on the rationality of performing the act that brings it about, not vice versa. Despite appearances, though, Anderson's view needn't be contrary to TCR-3. Whether it is or not depends on whether there is some further end,  $e_2$ , in virtue of which it is rational to perform the act that brings about  $e_1$ . For instance, it may be that my wanting and pursuing the state of affairs in which my wife and I spend time alone together is rational only because it is rational for me to want and pursue the state of affairs in which my actions adequately express my rational evaluative attitudes toward her. If this is right, then there is no conflict between Anderson's view and TCR-3, for if this is right, then what ultimately accounts for the rationality of performing a given action is the rationality of wanting some end to obtain, not vice versa. The issue, then, is whether or not Anderson's injunction to act only in ways that express one's rational attitudes toward persons, animals, and things is extensionally equivalent to the injunction to act so as to promote some end—i.e., some state of affairs.

In fact, Anderson argues that her injunction, which she calls "E," to act only in ways that *express* one's rational attitudes toward persons is not extensionally equivalent to the injunction, which she calls "P," to act so as to *promote* the state of affairs where one's actions adequately express one's rational attitudes. She argues that E is no more extensionally equivalent to P than the injunction, E' (i.e., E prime), to make only logically valid inferences is extensionally equivalent to the injunction, P', to act so as to promote the state of affairs where one makes logically valid inferences. In

both cases, the latter (P or P') tells one to do something that the former (E or E') does not: in the case of P, to perform more acts that adequately express one's rational attitudes toward persons and, in the case of P', to make more valid inferences. Moreover, unlike E, "P tells me to violate E, if by doing so I can bring about more events containing E than if I never violated E" (ANDERSON 1996, p. 544). Thus, unlike E, P tells me to betray my wife by committing adultery if my doing so will create an opportunity for me to more fully express my love for her than I otherwise would have, as where this betrayal will enable a reconciliation that allows me to express my love for her more fully than before (ANDERSON 1996, p. 545). Anderson shows, therefore, that E and P yield different prescriptions. But these differences in what E and P prescribe do not demonstrate that E is not equivalent to *some* injunction to promote a state of affairs; they only demonstrate that E is not extensionally equivalent to P.

So consider P\*: act so as to *promote* the state of affairs in which as few of one's actions as possible fail to adequately express one's rational attitudes. Unlike P, P\* does not tell one to act so as to promote greater instances of acting in ways that express one's rational attitudes toward persons. And, like E, it functions as a constraint, telling one not to violate E (i.e., not to act in a way that fails to adequately express one's rational attitudes toward persons) even in order to bring about more events containing E (i.e., more events where one adequately expresses one's rational attitudes toward persons). Of course, P\* directs an agent to violate E if her doing so will minimize her own violations of E. But, as Nicholas Sturgeon (1996, p. 521) has pointed out, it is hard to imagine any but the most fanciful of cases in which an agent's violating E now will prevent her from committing more numerous, comparable violations of E in the future. Nor is it obvious that, in such fanciful cases, it is E rather than P\* that gets the more intuitively plausible results.<sup>38</sup> But rather than debate the issue as to whether E or P\* is more plausible, it is clear that E is extensionally equivalent to P\*\*: act so as to *promote* the state of affairs in which one's current actions do not fail to adequately express one's rational attitudes toward persons.<sup>39</sup> P\*\* and E yield the exact same prescriptions, and so we see that there is nothing that E can account for that P\*\* cannot account for. And since the proponent of TCR-3 can adopt P\*\* and thereby account for

<sup>38</sup> For more on this issue, see PORTMORE 1998 and Chapter 4.

<sup>39</sup> For any deontological injunction prohibiting agents from performing a certain type of action, there is, as I and others have argued, an equivalent consequentialist injunction requiring agents to bring about a certain state of affairs. See, for instance, PORTMORE 2007 and PORTMORE 2008 and LOUISE 2004.

all the intuitive judgments that Anderson appeals to, there is no reason for us to think that Anderson's view, even if correct, is incompatible with TCR-3.

I've argued, then, that in many instances what an agent has reason to do depends on what she has reason to desire, not vice versa. For instance, it seems that I have reason to put my money in the higher-yielding savings account, because I have reason to prefer the state of affairs in which I earn more interest on my savings. And I've shown that, contrary to initial appearances, Anderson's view, even if correct, gives us no reason to deny this.

This brings us to the second alternative to TCR-3, which is to claim that what explains the truth of the bi-conditional 'PER if and only if DES' is that, in some instances, 'PER, because DES', in other instances, 'DES, because PER', but, in every instance, one or the other. The problem with this view is that it is fragmented, and thus its proponents owe us some explanation as to why sometimes reasons for acting are explanatorily prior and other times reasons for desiring are explanatorily prior. Until such an explanation is forthcoming, we should, I believe, accept the more unified and systematic TCR-3 instead.

The third and final alternative to TCR-3 is the view that what explains the truth of the bi-conditional 'PER if and only if DES' is that, in every instance, both 'PER, because BET' and 'DES, because BET'. We should, I think, reject this third alternative. Consider what BET might be. That is, consider what might possibly explain both PER and DES? The most obvious candidate is that 'BET' stands for ' $o_i$  is *better* than  $o_j$ '. But, if 'BET' stands for ' $o_i$  is *better* than  $o_j$ ', then we should reject the claim that 'DES, because BET'. As many philosophers have come to think, this just gets the order of explanation wrong, for to say that  $o_i$ 's obtaining would be good is just to say that  $o_i$  has the purely formal, higher-order property of having other properties that provide sufficiently weighty reasons of the right kind for desiring that  $o_i$  obtains.<sup>40</sup>

But even if we don't accept this, the buck-passing account of value, we should still reject the claim that, in every instance, 'DES, because BET', for it seems that sometimes agents have more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains even though  $o_i$  is *not* better than  $o_j$ . And so it can't be that, in every instance, what explains DES is BET. For

<sup>40</sup> Again, this is called the fitting-attitude or buck-passing account of value—see, for instance, RABINOWICZ AND RØNNOW-RASMUSSEN 2004, SCANLON 1998, and STRATTON-LAKE AND HOOKER 2006.

instance, it seems that, in the case from §§3.4.1, I have more reason to desire  $o_1$  (the possible world in which it's true to say that I ran as fast as I could and rescued my daughter) than I have to desire  $o_2$  (the possible world in which it's true to say that I refrained from running as fast as I could and thereby allowed Richard to rescue his two daughters) even though  $o_2$  is, impersonally speaking, better than  $o_1$ . In this case, it seems that my agent-relative reasons for preferring that my own daughter is saved outweigh my agent-neutral reasons for preferring the better outcome. Since, in this case, I have reason to prefer the worse outcome (viz.,  $o_1$ ), it can't be that what explains why I have more reason to desire that  $o_1$  obtains than to desire that  $o_2$  obtains is that  $o_1$  is better than  $o_2$ , for that's simply false.

So, if the only likely candidate for 'BET' is ' $o_i$  is *better* than  $o_j$ ', and if, on this interpretation of 'BET', we should reject 'DES, because BET', as I've argued, then we should reject this third alternative to TCR-3. Of course, these might seem like two big ifs. So the argument here is only presumptive. Given the prima facie plausibility of these antecedents, there is a presumptive case to be made against this third alternative to TCR-3. But even if the case is only a presumptive one, it's important to note that the argument should have persuasive force against some of TCR's leading critics, for critics such as Scanlon accept my assumption that it is DES that explains BET, not vice versa.

§§3.4.4 *In defense of TCR on the whole:* Admittedly, the arguments for TCR-1, TCR-2, and TCR-3 are each far from decisive. For one thing, the argument for TCR-1 and TCR-2 relied on certain intuitions that not everyone will share. For another, the argument for TCR-3 relied, in part, on the buck-passing account of value. But, again, it is worth noting that at least some of TCR's leading critics endorse the buck-passing account of value. Of course, my aim is not just to convince TCR's leading critics, but to convince others as well, and yet not everyone will both share my intuitions and accept my assumptions. So let me explain why I think that even those who don't, nevertheless, have good reason to accept TCR.

Compared to its alternatives, TCR does a far superior job of systematizing our various substantive convictions about what we have reason to do. Everyone seems to admit that in at least some instances an agent's reasons for performing a given action derive from her reasons for desiring its outcome. For instance, we think that, other things being equal, you have more reason to choose the more pleasure-inducing meal option to the less pleasure-inducing meal option if and only if, and because, you have more reason to desire the outcome in which you have experienced

more pleasure than to desire the outcome in which you have experienced less pleasure. No one thinks that it's the other way around: that you have more reason to want the outcome in which you have experienced more pleasure because you have more reason to choose the more pleasure-inducing meal option.

The only point in contention, then, is whether reasons for action always derive from reasons for desiring. Some have thought not, for they have thought that there are certain substantive views about value and practical reasons that TCR cannot accommodate—e.g., the view that it is certain concrete entities, and not states of affairs, that are the primary bearers of intrinsic value and that, when playing to win at a game of tennis, there are reasons not to take certain other considerations into account when deciding whether to make a certain strategic shot. But, as I've shown, this is a misconception: there are no substantive views about values and reasons that TCR cannot accommodate. TCR is an extremely ecumenical and accommodating view. Thus, another reason to accept TCR, apart from both its *prima facie* plausibility and the reasons there are to accept each of its three constitutive claims, is that TCR provides a systematic account of our various substantive views about reasons for action.

#### §CR

In this chapter, I've defended TCR against putative counterexamples, and I have tried to ward off many actual and potential misconceptions. I have shown that TCR is quite ecumenical and that, consequently, we do not have to give up any of our considered convictions about values or reasons to accept it. I have also offered positive arguments in favor of each of TCR's three claims. Moreover, I have argued that TCR is unsurpassed in its ability to systematize our considered convictions about practical reasons.

If the arguments in this and the preceding chapter succeed, we should, then, accept both MR and TCR. And these two theses entail a third: viz., act-consequentialism (or 'AC' for short). The argument, you'll recall, goes as follows:

- MR** An act's deontic status is determined by the agent's reasons for and against performing it, such that, if a subject, *S*, is morally required to perform an act, *x*, then *S* has most reason to perform *x*.
- TCR** The agent's reasons for and against performing an act are determined by her reasons for and against preferring its

outcome to those of the available alternatives, such that, if S has most reason to perform  $x$ , then, of all the outcomes that S could bring about, S has most reason to desire that  $x$ 's outcome obtains.

Therefore,

**AC** An act's deontic status is determined by the agent's reasons for and against preferring its outcome to those of the available alternatives, such that, if S is morally required to perform  $x$ , then, of all the outcomes that S could bring about, S has most reason to desire that  $x$ 's outcome obtains.<sup>41</sup>

But although we should accept AC, we should, as I argued in the preceding chapter, reject all traditional versions of AC, as all traditional versions of AC conflict with moral rationalism (MR). Thus, in the remaining chapters, I will show how we might go about constructing a version of consequentialism that doesn't conflict with MR—one that comports with our considered convictions both about what we are morally required to do and about what we are rationally required to do. I'll start, in the next chapter, by showing how any plausible nonconsequentialist theory can be consequentialized, which is to say that, for any plausible nonconsequentialist theory, we can construct a consequentialist theory that yields the exact same set of deontic verdicts that it yields. This will establish that we can consequentialize a nonconsequentialist theory whose deontic verdicts are more in tune with our considered moral convictions than any traditional version of consequentialism is.

<sup>41</sup> As I noted in the previous chapter, some might prefer to call this *rational-desire teleology*, since, unlike traditional act-consequentialism, it takes the deontic statuses of actions to be a function of the agent's reasons for desiring their outcomes as opposed to the value of their outcomes. Nevertheless, 'act-consequentialism' is term of art, which I will use it to refer to AC. The reader should keep in mind, though, that it doesn't matter what we call it. What matters is only whether AC is a plausible view.



## CHAPTER 4

---

# Consequentializing Commonsense Morality<sup>1</sup>

- 4.1 How to consequentialize
- 4.2 The deontic equivalence thesis
- 4.3 Beyond the deontic equivalence thesis: How consequentialist theories can do a better job of accounting for our considered moral convictions than even some nonconsequentialist theories can
- 4.4 The implications of the deontic equivalence thesis
- 4.5 An objection

UNLIKE UTILITARIANISM, NONCONSEQUENTIALIST theories—at least, the plausible ones<sup>2</sup>—tend to comport with our considered moral convictions. In so doing, plausible nonconsequentialist theories avoid the sorts of deontic verdicts that got utilitarianism in so much hot water vis-à-vis moral rationalism (MR). Recall from Chapter 2 that utilitarianism conflicts with MR, for it requires that agents maximize aggregate utility even when they have sufficient reason to do otherwise. Unlike utilitarianism, though, plausible versions of nonconsequentialism comport with MR—at least, they do given certain plausible assumptions about what agents have decisive reason to do.

In this chapter, I hope to demonstrate that any plausible nonconsequentialist theory can be consequentialized, which is to say that, for any plausible nonconsequentialist theory, we can construct a consequentialist theory that yields, in every possible world, the exact same set of deontic verdicts that it yields. Since plausible nonconsequentialist theories tend to yield deontic verdicts that comport with MR, this will establish that some consequentialist theories (specifically, those that are extensionally equivalent to their plausible nonconsequentialist

<sup>1</sup> Much of the material in this chapter is drawn from PORTMORE 2009b.

<sup>2</sup> Here, I have in mind those nonconsequentialist theories that have attained some prominence in the literature, theories such as Kantianism, Rossian pluralism (ROSS 1930), Scanlonian contractualism (SCANLON 1998), Hooker's rule-consequentialism (HOOKER 2000), neo-Aristotelian virtue ethics (FOOT 2001 and HURSTHOUSE 1999), Slote's agent-based virtue ethics (SLOTE 2001), Kamm's rights-based deontology (KAMM 2007), etc. (Recall that I use the term 'nonconsequentialism' as shorthand for '*non-act-consequentialism*'. Thus, Hooker's rule-consequentialism counts as a nonconsequentialist theory on my terminology.)

counterparts) comport with MR—again, given certain plausible assumptions about what agents have decisive reason to do. What’s more, these consequentialist theories will retain what we found most compelling about utilitarianism in the first place: its consequentialism, which we were driven to accept given the attractiveness of both MR and the teleological conception of reasons (TCR)—or so I argued in the last two chapters. Thus, by consequentializing some plausible nonconsequentialist theory, we can construct a substantive version of consequentialism that takes what is best from both utilitarianism and nonconsequentialism. From utilitarianism, this newly constructed version of consequentialism takes utilitarian’s compelling idea—viz., the idea that an act’s deontic status is determined by the agent’s reasons for preferring its outcome to those of the available alternatives such that it can never be morally wrong for her to act so as to bring about the outcome that she has most reason to want to obtain. And, from nonconsequentialism, it takes nonconsequentialism’s intuitively plausible deontic verdicts and their compatibility with MR.

The chapter has the following structure. In §4.1, I show how to consequentialize a nonconsequentialist theory. In §4.2, I demonstrate that, for any remotely plausible nonconsequentialist theory, there is a consequentialist counterpart theory that is deontically equivalent to it such that the two theories will be extensionally equivalent with respect to their deontic verdicts. I call this the *deontic equivalence thesis* or ‘DET’ for short. I go beyond previous attempts to demonstrate DET by showing that, for any remotely plausible nonconsequentialist theory, we can construct a consequentialist counterpart theory that mimics its verdicts not only about what’s permissible but also about what’s supererogatory. What’s more, I show that consequentialism can accommodate moral dilemmas, agent-centered options, agent-centered restrictions, and the self-other asymmetry.

In §4.3, I show that whereas consequentialism can accommodate any deontic verdict that nonconsequentialism can accommodate, certain prominent nonconsequentialist theories are unable to accommodate certain plausible deontic verdicts that consequentialism can accommodate. Thus, even just in terms of accommodating our considered moral convictions, consequentialism actually has a leg up on some of its nonconsequentialist rivals. In §4.4, I take issue with those philosophers who have claimed that one implication of DET is that every nonconsequentialist theory can be fairly regarded as a form of consequentialism, and that, therefore, we are all consequentialists. Lastly, in §4.5, I argue that although the consequentializer will need to appeal to our considered moral convictions

in determining how outcomes rank, this in no way renders the resulting consequentialist position circular or uninformative.

#### §4.1 How to consequentialize

The recipe for consequentializing a nonconsequentialist theory is quite simple: Take whatever considerations that the nonconsequentialist theory takes to determine the deontic statuses of actions and insist that those considerations determine how their outcomes rank. In this way, the consequentialist can produce a ranking of outcomes that when combined with her criterion of rightness yields, in every possible world, the same set of deontic verdicts that the nonconsequentialist theory yields, such that, for any deontic predicate ('permissible', 'impermissible', or 'obligatory'), the resulting consequentialist counterpart theory and the original nonconsequentialist theory will be in perfect agreement as to the set of actions that are in the extension of that predicate. Such theories are *deontically equivalent*.

To illustrate, suppose that we want to consequentialize some nonconsequentialist theory, turning it into a version of maximizing consequentialism, according to which an act is morally permissible if and only if its outcome is not outranked by that of any other available act alternative. The procedure for doing so is as follows. For any act that the nonconsequentialist theory takes to be morally permissible, deny that its outcome is outranked by that of any other available act alternative.<sup>3</sup> For any act that the nonconsequentialist theory takes to be morally impermissible, hold that its outcome is outranked by that of some available act alternative—specifically, outranked by all the outcomes of all permissible acts. And, for any act that the nonconsequentialist theory takes to be morally obligatory, hold that its outcome outranks that of every other available act alternative.<sup>4</sup> This procedure produces a ranking of outcomes that when combined with maximizing consequentialism's criterion of rightness yields a deontically equivalent version of maximizing consequentialism.

Of course, the rankings of outcomes will need to be agent-relative (and, perhaps, also time-relative) in order to consequentialize certain nonconsequentialist theories, such as those that incorporate agent-centered

<sup>3</sup> An act alternative is any of the mutually exclusive and jointly exhaustive act-tokens that are available to a given agent at a given time.

<sup>4</sup> I'm assuming that an act is morally obligatory if and only if it is the only morally permissible act alternative available to the agent. See VALLENTYNE 1989, section 3.

restrictions. To see this, suppose that I have the choice of performing either  $A_1$  or  $A_2$  (with corresponding outcomes  $O_1$  and  $O_2$ ) and that you have the choice of performing either  $A_3$  or  $A_4$  (with corresponding outcomes  $O_3$  and  $O_4$ ). Assume both that you will perform  $A_3$  if and only if I perform  $A_1$  and that you will perform  $A_4$  if and only if I perform  $A_2$ .<sup>5</sup> Assume that  $A_1$  involves my staying still, which will cause no one's death, and that  $A_2$  involves my moving about, which will cause the death of Alex. Assume that  $A_3$  involves your moving about, which will cause the death of both Alex and Bonnie, and that  $A_4$  involves your staying still, which will cause no one's death. So  $O_1$  is the possible world in which I stay still and you move about, thereby causing the death of both Alex and Bonnie. And  $O_2$  is the possible world in which I move about, thereby causing the death of Alex, and you stay still. And given that you will move about (i.e., perform  $A_3$ ) if and only if I stay still (i.e., perform  $A_1$ ) and that you will stay still (i.e., perform  $A_4$ ) if and only if I move about (i.e., perform  $A_2$ ), it follows that  $O_3$  is identical to  $O_1$  and that  $O_4$  is identical to  $O_2$ .

Let's suppose that the nonconsequentialist theory to be consequentialized holds that I'm obligated to perform  $A_1$  and that you're obligated to perform  $A_4$ .<sup>6</sup> In order for the consequentializer to hold that I'm obligated to perform  $A_1$ , she must hold that  $O_1$  outranks  $O_2$ .<sup>7</sup> And this implies that  $O_3$  outranks  $O_4$ , for  $O_3$  is identical to  $O_1$  and  $O_4$  is identical to  $O_2$ . Yet, for the consequentializer to hold that you're obligated to perform  $A_4$ , she must hold that  $O_4$  outranks  $O_3$ . How, then, can the consequentializer have it both ways: that  $O_4$  both outranks, and is outranked by,  $O_3$ ? The answer is that the rankings will need to be agent-relative. On *my* ranking,  $O_1$  outranks  $O_2$ , which implies that, on *my* ranking,  $O_3$  outranks  $O_4$ . Nevertheless, on *your* ranking,  $O_4$  outranks  $O_3$ , which is consistent with  $O_3$  outranking  $O_4$  on *my* ranking. Thus, the maximizing consequentialist must hold that an agent is morally permitted to perform a given act if and only if its outcome is not, *on the agent's ranking*, outranked by that of any other available act alternative.<sup>8</sup>

<sup>5</sup> Recall that I'm assuming that counterfactual determinism is true—that for each act alternative available to the agent there is some determinate fact as to what the world will be like if the agent performs the act. See note 1 of Chapter 3.

<sup>6</sup> I'm assuming that although you *will* perform  $A_3$  if I perform  $A_1$ , you *can* perform  $A_4$  even if I perform  $A_1$ .

<sup>7</sup>  $O_1$  and  $O_2$  actually represent sets of outcomes. This complicates things a bit. Let us say that  $O_1$  outranks  $O_2$  if and only if, for each outcome in  $O_2$ , the outcome in  $O_1$  that's most similar to it outranks it.

<sup>8</sup>  $S$ 's evaluative ranking is not a ranking in terms of what  $S$  subjectively prefers, but rather a ranking in terms of what  $S$  ought, objectively speaking, to prefer.

## §4.2 The deontic equivalence thesis

I've claimed that, for any remotely plausible nonconsequentialist theory, we can construct a version of consequentialism that is deontically equivalent to it. This, again, is the *deontic equivalence thesis*.<sup>9</sup> Rather than take a piecemeal approach to demonstrating the deontic equivalence thesis, I'll attempt a more systematic approach, showing how consequentialism can accommodate all the various general features that can be found in various nonconsequentialist theories: moral dilemmas, agent-centered options, supererogatory acts, the self-other asymmetry, agent-centered restrictions, and special obligations. These are the features of commonsense morality.<sup>10</sup> If I can show that all these features can be consequentialized, then it should be easy enough to see how we could consequentialize the various deontic verdicts that are instances of these. I'll proceed in the above order.

§§4.2.1 *Moral dilemmas*: As Peter Vallentyne notes, a "moral dilemma arises when an agent is in a choice situation in which he/she cannot satisfy the dictates of morality" (1989, 301). Various nonconsequentialist theories allow for moral dilemmas. Take, for instance, a deontological theory that includes an absolute prohibition against breaking one's promises. On such a theory, a moral dilemma will arise whenever an agent makes a set of promises that cannot all be kept. Suppose, for instance, that, at  $t_1$ , I promised Cristobal that I would perform  $A_1$  at  $t_5$ . Further suppose that, due to a lapse in memory, I promised Dolly at  $t_3$  that I would not perform  $A_1$  at  $t_5$ . It's now  $t_4$ , and I must decide whether or not to perform  $A_1$  at  $t_5$ . Promise-breakings are, we're supposing, absolutely prohibited, but I cannot both perform  $A_1$  and not perform  $A_1$  at  $t_5$ . So, no matter what, I will break one of my two promises and thereby act impermissibly.<sup>11</sup> I face, therefore, a moral dilemma.

<sup>9</sup> This thesis sometimes goes by other names, such as the "Representation Thesis" (ODDIE & MILNE 1991) and "Dreier's Conjecture" (BROWN 2004). The thesis, or some close cousin to it, has been defended and/or endorsed by Brown (2004), Dreier (1993), Dreier (2011), Louise (2004), Oddie & Milne (1991), Portmore (2007a and 2009b), Smith (2009), and Vallentyne (1988).

<sup>10</sup> Actually, I'm not confident that commonsense morality—the morality of our commonsense moral convictions—includes moral dilemmas, but I'll consider moral dilemmas in any case.

<sup>11</sup> Thus, on this theory, the common assumption that there is always at least one permissible act alternative available to the agent is false. And if this common assumption is false, then it's also false to suppose that S is obligated to perform  $x$  if and only if it would be wrong for

Although some philosophers deny that such moral dilemmas are conceptually possible, there are many who think that they are. Nonconsequentialist theories that include such moral dilemmas are, therefore, at least remotely plausible. So if the deontic equivalence thesis is true, then consequentialism must be able to accommodate such moral dilemmas. I will argue that there are two types of moral dilemmas (prohibition dilemmas and obligation dilemmas), that consequentialism can accommodate only one of the two (i.e., prohibition dilemmas), but that the other (i.e., obligation dilemmas) is clearly conceptually impossible.<sup>12</sup>

First, let's consider obligation dilemmas. These are choice situations in which more than one of an agent's available act alternatives is obligatory. An example would be where an agent is both morally obligated to perform  $A_1$  and morally obligated to perform  $A_2$ , can perform each of  $A_1$  and  $A_2$ , but cannot perform both  $A_1$  and  $A_2$ . And for this to be a genuine instance of an obligation dilemma,  $A_1$  and  $A_2$  must each be a *positive* act—i.e., not the "act" of refraining from performing some act.

Consequentialism cannot accommodate such dilemmas. Here's why. On (maximizing) consequentialism, an act is obligatory only if its outcome outranks (on the agent's ranking) those of its alternatives. So  $A_1$  is morally obligatory only if  $O_1$  outranks  $O_2$ . And  $A_2$  is morally obligatory only if  $O_2$  outranks  $O_1$ . So in order for both  $A_1$  and  $A_2$  to be obligatory,  $O_1$  would have to both outrank, and be outranked by,  $O_2$  on the same ranking.<sup>13</sup> And that's not possible.<sup>14</sup> So it's impossible for consequentialism to accommodate obligation dilemmas. But this is not at all surprising, since obligation dilemmas are conceptually impossible. To see this, note that an act is obligatory if and only if both (i) it is permissible and (ii) all its

S to refrain from performing  $x$ , for, if this common assumption is false, then it could be wrong for S to perform  $x$  even though it would also be wrong for S to refrain from performing  $x$ , and S cannot be obligated to do what's wrong. We should, then, deny that S is obligated to perform  $x$  if and only if it would be wrong for S to perform any available alternative and accept instead that S is obligated to perform  $x$  if and only if *both* (i) S is permitted to perform  $x$  and (ii) it would be wrong for S to perform any available alternative (VALLENTYNE 1989, p. 305).

<sup>12</sup> The distinction between obligation dilemmas and prohibition dilemmas as well as the forthcoming argument for obligations dilemmas being conceptually impossible come from VALLENTYNE 1989.

<sup>13</sup> It has to be the same ranking, because there is only one agent involved, who is deciding whether to perform  $A_1$  or  $A_2$  at some specific time. Thus, it can't be that we have two different agent-relative or time-relative rankings.

<sup>14</sup> I'm assuming here that the  $x$ -outranks- $y$  relation is asymmetric. A relation, R, is asymmetric just in case: if  $xRy$ , then  $\sim yRx$ .

available alternatives are impermissible (VALLENTYNE 1989, p. 305).<sup>15</sup> From (ii), it follows that, if  $A_1$  is obligatory, then  $A_2$ , which is an available alternative, must be impermissible. But if  $A_2$  is obligatory, then, according to (i),  $A_2$  must be permissible. So, for both  $A_1$  and  $A_2$  to be obligatory,  $A_2$  would have to be both permissible and impermissible. And that's not logically possible. Let's turn, then, to those moral dilemmas that are conceptually possible: namely, prohibition dilemmas.

A prohibition dilemma is a choice situation in which all the available act alternatives are impermissible. Can consequentialism accommodate such dilemmas? Yes. For instance, there will be such dilemmas whenever an agent faces a choice situation in which there are an infinite number of successively higher-ranked outcomes that she can bring about. To illustrate, suppose that an agent faces a choice situation in which God will make the total aggregate utility be whatever natural number she picks and that, if she fails to pick a natural number, God will make the total aggregate utility zero. No matter what she does, then, she acts wrongly. If she picks no natural number, she acts wrongly. And if she picks some natural number,  $n$ , she acts wrongly, for had she picked  $n+1$  instead, there would have been even more aggregate utility.<sup>16</sup>

Of course, this establishes only that consequentialism can accommodate some prohibition dilemmas. To establish the deontic equivalence thesis, though, I need to show not only that consequentialism can accommodate prohibition dilemmas, but also that consequentialism can accommodate the various sorts of prohibition dilemmas that any remotely plausible nonconsequentialist theory might entail. So, for instance, the consequentialist needs to be able to accommodate dilemmas of the sort that I considered at the beginning of this subsection, where, for instance, I will act wrongly if I break my promise to Cristobal by performing  $A_1$  at  $t_5$  and where I will act wrongly if I break my promise to Dolly by not performing  $A_1$  at  $t_5$ .<sup>17</sup> To illustrate, then, let  $O_1$  be the set of all the outcomes

<sup>15</sup> See note 11 for why the proponent of moral dilemmas cannot instead hold that  $S$  is obligated to perform  $x$  if and only if it would be wrong for  $S$  to refrain from performing  $x$ . And see VALLENTYNE 1989 (pp. 307-9) for a defense of the definition 'obligatory' given in the body of this paper.

<sup>16</sup> Natural numbers are all finite.

<sup>17</sup> It may seem that the existence of a prohibition dilemma entails the existence of an obligation dilemma, for a prohibition against performing  $x$  is just an obligation not to perform  $x$ . In actuality, though, the existence of a prohibition dilemma does not entail the existence of an obligation dilemma. This is because obligation dilemmas, by definition, consist in obligations to perform two or more mutually exclusive *positive* acts, whereas the obligation not to perform  $x$  does not constitute an obligation to perform any positive act. To

corresponding to the acts in  $A_1$ .<sup>18</sup> And let  $A_2$  be the set of all the (positive) act-tokens,  $a_2, a_3, \dots, a_n$ , that entail my not performing  $A_1$  at  $t_5$ , and let  $O_2$  be the set of all their corresponding outcomes. And let's assume that I'm conscious and deliberating at  $t_4$  and so must perform some positive act at  $t_5$ .<sup>19</sup> According to the nonconsequentialist theory to be consequentialized, it is both wrong for me to perform  $A_1$  and wrong for me to perform  $A_2$ , and yet I must do one or the other, for  $A_1$  and  $A_2$  are jointly exhaustive. No matter what I do, then, I will act wrongly. How can the consequentialist accommodate such prohibition dilemmas?

To do so, the consequentialist must hold that whether or not  $O_1$  outranks  $O_2$  depends on whether I perform  $A_1$  or  $A_2$ .<sup>20</sup> More specifically,

illustrate, suppose that I promised, at  $t_2$ , not to perform  $A_1$  at  $t_5$  but then died at  $t_4$ . In this case, I managed not to violate my obligation not to perform  $A_1$  at  $t_5$  even though I didn't perform any positive act at  $t_5$ . Thus, the obligation not to perform  $x$  does not consist in an obligation to perform any positive act. And this means that the situation in which I've promised both to perform  $A_1$  at  $t_5$  and not to perform  $A_1$  at  $t_5$  constitutes a prohibition dilemma, but not an obligation dilemma.

<sup>18</sup> Let me just stipulate that there is only a finite set of outcomes available to me in this choice situation. Thus, it will not be possible for the consequentializer to employ the original strategy for accommodating prohibition dilemmas, which was to postulate that, for each available outcome, there is always some other available outcome that outranks it.

<sup>19</sup> I'm assuming that intentionally remaining motionless is itself a positive act. One performs a positive act when one intentionally does something, whether it be, say, walking, running, or remaining motionless. It's possible for S to not perform a positive act,  $x$ , without performing any other positive act instead. For instance, if S is unconscious at  $t_5$ , then S is not performing  $x$  at  $t_5$ , but neither is S performing some other positive act at  $t_5$ .

<sup>20</sup> This might sound strange, but Campbell Brown has offered the following example to illustrate this possibility:

Suppose...that we endorse a theory according to which agents ought always to act so as to maximise the aggregate preference satisfaction of all people, ...[and] we quite explicitly insist that only the preferences of *actual* people are to be taken into account.... Now, consider the following case. A couple would like to have a child; if they were to have a child, this would increase their preference satisfaction. However, the child would prefer that she didn't exist. Moreover, the child's preference would be sufficiently strong or intense as to outweigh the preferences of the parents.... What ought the parents to do, according to our theory? Suppose first that they do have the child. In this case the child is an actual person, and so her preferences count in determining what the parents ought to do. And, since the child's preference outweighs the parents' preferences, our theory implies that the parents acted wrongly; they ought not to have had the child. Suppose next that the parents don't have the child. In this case, the child is a merely possible person, and so her "preferences" don't count. Thus, again, our theory implies that the parents acted wrongly. (BROWN 2004, p. 33)

The idea that consequentialists can accommodate moral dilemmas in this way is Brown's.



the consequentialist must hold that  $O_1$  outranks  $O_2$  if and only if I perform  $A_2$  and that  $O_2$  outranks  $O_1$  if and only if I perform  $A_1$ . The idea is that no matter which I perform, I should regret performing the act that I did, and, thus, I should wish that the other act's outcome obtains instead. So if I perform  $A_1$ , I should wish/prefer that  $O_2$  obtains, but if I perform  $A_2$ , I should wish/prefer that  $O_1$  obtains. This ensures that no matter which act I perform, its outcome will be outranked, implying that the act is impermissible. Thus, the consequentialist can, given certain assumptions, accommodate such prohibition dilemmas.

§§4.2.2 *Agent-centered options, supererogatory acts, and the self-other asymmetry*: A supererogatory act is, as the name suggests, an act that goes above and beyond what's morally required. Clearly, satisficing consequentialism can accommodate such acts. On satisficing consequentialism, agents are morally required only to produce an outcome that's ranked sufficiently high. Producing an outcome that ranks even higher is supererogatory. But since it is doubtful that satisficing consequentialism is even tenable, the consequentializer should hope that some non-satisficing version of consequentialism, such as maximizing consequentialism, can also accommodate supererogatory acts.<sup>21</sup> Yet, as Michael Byron has pointed out, "It's difficult to see how a maximizing conception of morality can allow room for supererogation: If I'm required in every case to choose the best available option, how could I ever do *more* than morality requires?" (BYRON 2004, p. 9) As we'll see below, the answer depends on what doing *more* than is required consists in.

To illustrate, suppose that we were to define a supererogatory act as one that involves a greater self-sacrifice for the sake of others than is required.<sup>22</sup> In that case, even maximizing act-utilitarianism will accommodate supererogatory acts, because sometimes an agent will have a choice between two or more maximizing acts that differ with respect to whether the maximal quantity of utility is distributed in greater proportion to others or in greater proportion to the agent. Such maximizing alternatives will be morally optional on act-utilitarianism, and those that involve *more* of the maximal quantity of utility being distributed to others as opposed to the agent will, on the above definition, be supererogatory.

<sup>21</sup> See BRADLEY 2006 for a refutation of satisficing consequentialism.

<sup>22</sup> Harwood (2003, p. 182) defines supererogation in this way and, consequently, concludes that maximizing act-utilitarianism can accommodate supererogatory acts. See also VESSEL 2010. See also MCNAMARA 1996a, p. 433.

But we might wonder whether this definition gets it right, specifically whether it correctly specifies the relevant sense of 'doing more than is required'. In specifying that 'doing more than is required' entails 'doing more for others than is required', this definition rules out the possibility of supererogation with respect to self-regarding duties. Yet it certainly seems possible to go above and beyond what such duties require.<sup>23</sup> For instance, we might think that there is a duty to develop one's talents and that this is an imperfect duty, and, thus, one that doesn't require that we take every opportunity to develop our talents, but requires only that we develop our talents to a certain extent, taking advantage of a sufficient number of opportunities to develop our talents. But if this is right, then it is surely possible to go above and beyond one's duty to develop one's talents, and yet doing so may be of no benefit to others. Here, then, is a plausible candidate for supererogation that the above definition simply rules out of court.<sup>24</sup>

And the definition faces another, potentially more serious, problem, for it seems that a supererogatory act must involve doing more of whatever there is *moral* reason to do more of.<sup>25</sup> For instance, we wouldn't think that perspiring more than is required is supererogatory, because we don't think that there is any moral reason to perspire more than is required. We don't think, for instance, that if there are two means to fulfilling some duty, then taking the means that involves more perspiration would, in virtue of that fact alone, be supererogatory. So we should think that doing more for others than is required is supererogatory only if (and because) there is some moral reason to do more for others than is required. So even if it's true that doing more for others than is required is supererogatory, this isn't fundamental to the definition of 'supererogatory'. Instead, what's fundamental is that an act is supererogatory only if there is more moral reason to perform it than to perform some other permissible alternative.<sup>26</sup>

<sup>23</sup> See KAWALL 2003, PORTMORE 2007a, and PORTMORE 2008a.

<sup>24</sup> If we think that supererogatory acts must involve self-sacrifice, then we need only imagine a case where developing one's talents beyond the extent to which one is required to develop them will involve self-sacrifice.

<sup>25</sup> Moral reasons are, of course, a proper subset of reasons for action. A moral reason is a reason that, morally speaking, counts in favor of, or against, performing some action. See PORTMORE 2008a and Chapter 5.

<sup>26</sup> I leave open the question of whether there are any further necessary conditions, such as: an act is supererogatory only if it requires more self-sacrifice (or is otherwise more difficult or demanding to perform) than some other permissible alternative that one has less moral reason to do.

If this is correct, then maximizing act-utilitarianism cannot accommodate supererogatory acts, for, on maximizing act-utilitarianism, the only time that one has the option of, say, performing either  $a_1$  or  $a_2$  is when there is no more moral reason to perform  $a_1$  than to perform  $a_2$  and vice versa. On maximizing act-utilitarianism, agents can never do more than they are required to do in the relevant sense of 'doing more': the sense that involves doing more of whatever one has moral reason to do, for, on maximizing act-utilitarianism, the only thing that an agent has moral reason to do is to promote utility. But since maximizing act-utilitarianism requires that agents maximize utility, there is never a situation in which an agent is permitted to do anything other than what she has most moral reason to do. Thus, there is, on maximizing act-utilitarianism, no room for supererogatory acts—acts that there is more moral reason to perform than to perform some other permissible alternative. It remains unclear, then, how a maximizing consequentialist theory might accommodate supererogatory acts.

As a matter of fact, the consequentialist can accommodate supererogatory acts. To do so, she need only adopt what I call *dual-ranking act-consequentialism*:

**DRAC** S's performing  $x$  is morally permissible if and only if, and because, there is no available act alternative that would produce an outcome that S has both more moral reason and more reason, all things considered, to want to obtain than to want  $x$ 's outcome to obtain.<sup>27</sup>

As we'll see in Chapter 7, things are actually quite a bit more complicated than I make them out to be here. As we'll see, I've been using the term 'supererogatory', whereas I should probably be using the term 'superperfecterogatory' instead. A superperfecterogatory act is one that goes above and beyond what perfect duty requires. The term 'supererogatory', by contrast, is best reserved for acts that go above and beyond what all duty (perfect and imperfect) requires. For now, though, I'll ignore this distinction and wait until Chapter 7 before taking up these more complicated issues.

<sup>27</sup> To the best of my knowledge, Ted Sider was the first to propose a version of DRAC, and he did so in order to accommodate the self-other asymmetry. His version of DRAC is called *Self/Other Utilitarianism* (SOU): S's performing  $x$  is morally permissible if and only if there is no available act alternative that would produce both more utility for others and more overall utility than  $x$  would—see SIDER 1993, p. 128.

I'm assuming that not only are there moral and non-moral reasons to perform certain actions, but also moral and non-moral reason to prefer certain outcomes to various others. It seems to me, for instance, that whereas the reason that you have to prefer an outcome in which you receive more pleasure to an otherwise identical outcome in which you receive less pleasure is a non-moral reason, the reason that you have to prefer an outcome in which

I call this *dual-ranking* act-consequentialism, because, on this view, outcomes are ranked according to two auxiliary rankings and the permissibility of actions is a function of a principal ranking that is in turn a function of these two auxiliary rankings. The two auxiliary rankings are AR<sub>1</sub> (a ranking of outcomes in terms of how much moral reason the agent has to want each of them to obtain) and AR<sub>2</sub> (a ranking of outcomes in terms of how much reason, all things considered, the agent has to want each of them to obtain). And one outcome is outranked by another on the principal ranking if and only if the one is outranked by the other on both of the auxiliary rankings. Because, on DRAC, any act whose outcome is not outranked by that of some alternative on both AR<sub>1</sub> and AR<sub>2</sub> is permissible, some permissible acts have outcomes that rank higher on AR<sub>1</sub> than those of other permissible alternatives. And since AR<sub>1</sub> is a ranking of outcomes in terms of how much moral reason the agent has to act so as to bring them about, these acts are supererogatory.<sup>28</sup> They are acts that the agent has more moral reason to perform than to perform some other permissible alternative.<sup>29</sup>

To illustrate, it will be helpful to consider a substantive version of DRAC—viz., *Schefflerian utilitarianism* (SU).<sup>30</sup> Consider its *criterion of rightness*:

children in the Third World don't starve to an otherwise identical outcome in which they do is a moral reason. Your lacking a preference for more personal pleasure would not entail a moral failing, but your lacking a preference for children not starving would. The idea that morality governs not only intentions to act, but also various other mental attitudes, such as desire, is not new. After all, we think that many moral sins involve having certain mental attitudes—e.g., lust, greed, envy, and pride.

<sup>28</sup> I'm assuming that if, as I argued in Chapter 3, how much reason, *all things considered*, an agent has to perform an act is determined by how much reason, *all things considered*, the agent has for preferring its outcome to those of the available alternatives, then how much *moral* reason an agent has to perform an act is determined by how much *moral* reason the agent has for preferring its outcome to those of the available alternatives.

<sup>29</sup> Adopting DRAC's dual-ranking structure may seem unmotivated, but I address this and other objections to DRAC in Chapter 5.

<sup>30</sup> Schefflerian utilitarianism is the view according to which all the following hold: (1) dual-ranking act-consequentialism is true; (2) S has more moral reason to want  $o_i$  to obtain than to want  $o_j$  ( $j \neq i$ ) to obtain if and only if  $o_i$  contains more utility for others (i.e., for those other than S) than  $o_j$  does; and (3) S has more reason, all things considered, to want  $o_i$  to obtain than to want  $o_j$  to obtain if and only if  $o_i$  contains more egoistically-adjusted utility than  $o_j$  does, where egoistically-adjusted utility includes everyone's utility but adjusts the overall total by giving S's utility, say, ten times the weight of anyone else's.

**SU<sub>α</sub>** S's performing  $x$  is morally permissible if and only if there is no available act alternative that would produce both more utility for others (i.e., for those other than S) and more egoistically-adjusted utility than  $x$  would, where egoistically-adjusted utility includes everyone's utility but adjusts the overall total by giving S's utility, say, ten times the weight of anyone else's.<sup>31</sup>

Unlike act-utilitarianism, SU accommodates both supererogatory acts and *agent-centered options*—options either to act so as to make things better overall but worse for oneself (or others) or to act so as to make things better for oneself (or others) but worse overall.<sup>32</sup> To illustrate, consider the scenario below, where an agent named Edouard has the following four mutually exclusive and jointly exhaustive options:  $a_1$ ,  $a_2$ ,  $a_3$ , and  $A_4$ . Here,  $A_4$  is an act-type—viz., that of performing something other than  $a_1$ ,  $a_2$ , or  $a_3$ . The consequences of each of these options are laid out in Table 4.1 below:<sup>33</sup>

**Table 4.1**

Let  $U_s(x)$  = the utility that accrues to S if S performs  $x$ , let  $U_{-s}(x)$  = the utility that accrues to others if S performs  $x$ , let  $U_{+s}(x) = U_{-s}(x) + [10 \times$

<sup>31</sup> I call this 'Schefflerian utilitarianism', because, like the theory that Scheffler (1994) argues for, it accommodates agent-centered options (or what Scheffler calls 'agent-centered prerogatives'), but eschews agent-centered constraints (or what Scheffler calls 'agent-centered restrictions'). SU does, however, differ from Scheffler's Distributive Hybrid Theory in at least two important respects: (1) it doesn't give priority to benefiting the worse off and (2) it incorporates not only agent-favoring options, but also agent-sacrificing options. (I explain the difference between the two below.)

Schefflerian utilitarianism is nearly identical to Jean-Paul Vessel's egoistically-adjusted self-other version of utilitarianism—see VESSEL 2010. I published my first formulations of Schefflerian utilitarianism and defended its ability to accommodate supererogatory acts in PORTMORE 2007a. Since then, I have adopted Vessel's apt phrase 'egoistically-adjusted' utility.

<sup>32</sup> Traditional forms of act-consequentialism allow for moral options whenever two or more acts maximize the good, but they never allow for agent-centered options, which permit agents either to act so as to make things better overall but worse for oneself (or others) or to act so as to make things better for oneself (or others) but worse overall. On traditional act-consequentialism, it is never morally permissible to make things worse overall; maximizing the overall good is, on this view, always obligatory.

<sup>33</sup> I've emphasized the fourth and fifth columns with bold-faced type since these two correspond to AR<sub>1</sub> and AR<sub>2</sub> of SU, respectively.

$U_s(x)$ ], and let  $U(x)$  = the overall utility that is produced if S performs  $x$ .

$x$	$U(x)$	$U_s(x)$	$\underline{U}_{-s}(x)$	$\underline{U}_{+s}(x)$	deontic status
$a_1$	17	7	10	80	merely permissible
$a_2$	6	2	4	24	impermissible
$a_3$	16	1	15	25	supererogatory
$A_4$	19	-1	20	10	supererogatory

SU accommodates agent-centered options in that it permits agents to give their own utility anywhere from no weight at all up to ten times the weight of any other. Thus, Edouard can permissibly choose to do what is better for himself and perform  $a_1$  as opposed to  $A_4$  even though performing the latter would be both better for others and better in terms of the impersonal good.<sup>34</sup> Edouard has, then, the *agent-favoring option* of performing  $a_1$  instead of  $A_4$ .<sup>35</sup> Furthermore, by allowing agents to give their own utility no weight, SU accommodates *agent-sacrificing options*, which permit agents to forgo their own greater good for the sake of providing others with some lesser net benefit.<sup>36</sup> For instance, Edouard can permissibly choose to perform  $a_3$  as opposed to  $a_1$ , thereby providing others with a net benefit of five utiles at a cost of six utiles to himself.<sup>37</sup>

Thus, SU can accommodate what's known as the *self-other asymmetry*: the asymmetry found in commonsense morality between what agents are permitted to do to themselves and what agents are permitted to do to others. It seems that although an agent is not permitted to be indifferent toward the welfare of others, she is permitted to be indifferent toward her own welfare. She can, for instance, sacrifice her own hand so as to save Fay's thumb even though it would be impermissible for her to sacrifice Gustav's hand to save Fay's thumb (even if Gustav consents to her doing so). On commonsense morality, then, agents are permitted to sacrifice their own greater good for the sake of providing others with some lesser net

<sup>34</sup>  $U(a_1)$  is 17, whereas  $U(A_4)$  is 19.

<sup>35</sup> Agent-favoring options are options either to promote one's own interests or to sacrifice one's own interests for the sake of doing more to promote the overall good.

<sup>36</sup> Agent-sacrificing options are options either to maximize the impersonal good or to sacrifice one's own self-interests for the sake of doing more to promote the interests of others even when doing so is not what's best overall, as where the extent of one's own self-sacrifices would exceed the extent of the benefits thereby bestowed upon others.

<sup>37</sup> The benefit to others is  $U_{-s}(a_3) - U_{-s}(a_1)$ , which equals 5. The cost to herself is  $U_s(a_3) - U_s(a_1)$ , which equals -6.

benefit, but they are not permitted to sacrifice someone else's greater good for the sake of providing others with some lesser net benefit—hence, the asymmetry between self and others.<sup>38</sup> Of course, agents are not always permitted to sacrifice their own welfare. It would, for instance, be wrong, as SU implies, for Edouard to choose to perform  $a_2$  as opposed to  $a_1$ . Such a choice would not only be worse for Edouard, but worse for others as well.<sup>39</sup>

SU can also accommodate supererogatory acts, or, at least, they can if the following is a plausible account of supererogation.

$S$ 's performing  $x$  is supererogatory if and only if there exists some available alternative,  $y$ , such that: (a)  $S$  is both morally permitted to perform  $x$  and morally permitted to perform  $y$  and (b)  $S$  has more moral reason to perform  $x$  than to perform  $y$ .<sup>40</sup>

Certain acts will, on SU, meet these conditions. SU is, as I've stipulated, a species of DRAC. Thus, like DRAC, SU has two auxiliary rankings: (AR<sub>1</sub>) a ranking of outcomes in terms of their utility for others, and (AR<sub>2</sub>) a ranking of outcomes in terms of their egoistically-adjusted utilities. This means that AR<sub>1</sub> of SU is a ranking of outcomes in terms of how much moral reason the agent has to bring them about. It specifies that the more utility for others an outcome contains the more moral reason there is to want it to obtain and to act so as to ensure that it does obtain. Thus, Edouard has more moral reason to perform  $a_3$  than to perform  $a_1$ , a permissible alternative, for  $a_3$  produces more utility for others. And the same holds for  $A_4$ .<sup>41</sup> Thus  $a_3$  and  $A_4$  would both be supererogatory on SU.

<sup>38</sup> Another way of putting it is that whereas the fact that  $S$ 's performing  $x$  would further  $S$ 's self-interest does not constitute a moral reason for  $S$  to perform  $x$ , the fact that  $S$ 's performing  $x$  would further someone else's self-interest does constitute a moral reason for  $S$  to perform  $x$ . For more on the self-other asymmetry, see SLOTE 1984, SIDER 1993, and SPLAWN 2001.

<sup>39</sup> Whereas  $U_s(a_1)$  is 7,  $U_s(a_2)$  is only 2. And whereas  $U_{-s}(a_1)$  is 10,  $U_{-s}(a_2)$  is only 4.

<sup>40</sup> Again, as we'll see in Chapter 7, these conditions best describe only those acts that go above and beyond what *perfect* duty requires. In Chapter 7, though, I'll give an account of what it means for acts to go above and beyond what duty (*perfect and imperfect*) requires. Also, I realize that some might want to add to conditions *a* and *b* a third condition: viz., (c)  $S$ 's performing  $x$  is more morally praiseworthy than  $S$ 's performing  $y$ . But I prefer to let the term 'supererogatory' denote the pre-theoretic notion of going above and beyond the call of duty (i.e., that of doing more than one has to), and an act can go beyond the call of duty without being praiseworthy, as where the agent performs the act with a bad motive—see MCNAMARA 2011. In letting the term 'supererogatory' refer to the pre-theoretical notion of exceeding the minimum that morality demands, I follow MCNAMARA 1996A.

<sup>41</sup>  $U_{-s}(a_3)$  and  $U_{-s}(A_4)$  are 15 and 20, respectively, whereas  $U_{-s}(a_1)$  is only 10.

What's more, SU allows for a range of supererogatory alternatives, where some supererogatory alternatives involve going further beyond what's required than others do.<sup>42</sup> For instance,  $A_4$  involves going further beyond the minimum that morality requires than  $a_3$  does.

§§4.2.3 *Agent-centered restrictions and special obligations*: Remarkably, we've just seen that a certain (non-satisficing) version of consequentialism, viz., SU (which is a species of DRAC), can accommodate many of the basic features of commonsense morality, including supererogation, agent-centered options, and the self-other asymmetry. But can consequentialism accommodate *agent-centered constraints* as well? An agent-centered constraint is a constraint on maximizing the good that it would be wrong to *infringe* upon even in some circumstances in which doing so would minimize comparable infringements of that constraint.<sup>43</sup> There are two types of agent-centered constraints: *agent-centered restrictions* and *special obligations*. Agent-centered restrictions *prohibit* agents from performing certain types of acts (e.g., murder) even in some circumstances in which doing so would minimize comparable commissions of that act-type. Special obligations, by contrast, *require* agents to do certain things (e.g., to keep their promises) even in some circumstances in which their failing to do so would minimize comparable failings of that type. These special obligations often arise as the result of performing certain past acts (such as, making a promise) but also come with occupying certain roles (such as, certain professional and familial roles).

To illustrate how we could fairly easily, although rather crudely, modify SU so as to accommodate both agent-centered restrictions and special obligations in addition to supererogatory acts, agent-centered

<sup>42</sup> On commonsense morality, there is often a range of supererogatory acts, where some supererogatory acts go further beyond the minimum required than others do (MCNAMARA 1996a). For instance, it seems that not only is my choosing to spend my Saturdays helping the poor supererogatory, but so is my choosing to spend both my Saturdays and my Sundays helping the poor. Although both choices would seem to be supererogatory, the latter seems to go further beyond the minimum required of me than the former does.

<sup>43</sup> If there is a constraint against performing a certain act-type, then any commission of an act of that type constitutes an *infringement* of that constraint. But not all infringements of a constraint are wrong. All and only those infringements that are wrong constitute *violations* of that constraint. I follow Thomson (1990, p. 122) and Zimmerman (1996, p. 181) in making this terminological distinction.

Constraints needn't be absolute. It may be permissible to infringe upon a constraint if enough good is at stake, and it may be permissible to infringe upon a constraint if doing so would prevent a sufficient number of comparable infringements of that constraint.



options, and the self-other asymmetry, consider *modified SU* and its *criterion of rightness*:

**Mod-SU<sub>c</sub>** S's performing  $x$  is morally permissible if and only if there is no available act alternative that would produce both more constraint-adjusted utility and more comprehensively-adjusted utility than  $x$  would, where the constraint-adjusted utility of an outcome is just the sum of the utility for others, adjusted by multiplying any disutility (or loss of utility) resulting from S's infringement of an agent-centered constraint by 500, and where the comprehensively-adjusted utility of an outcome is just its constraint-adjusted utility added to the product of S's utility times ten.

Mod-SU could easily accommodate a non-absolute constraint against, say, committing murder.<sup>44</sup> Agents would, other things being equal, be prohibited from committing murder even so as to prevent up to 500 others from each committing some comparable murder. And Mod-SU could easily accommodate a non-absolute constraint against failing to fulfill the special obligation that parents have to nurture their own children. Parents would, other things being equal, be prohibited from failing to nurture their own children even so as to prevent up to 500 others from each comparably failing to nurture their children. Clearly, then, Mod-SU can accommodate both agent-centered restrictions and special obligations.

Of course, Mod-SU is terribly crude. But I offer Mod-SU, not as a plausible moral theory, but only as a simple illustration of how a species of DRAC could accommodate both agent-centered restrictions and special obligations. More sophisticated versions of DRAC would adopt more sophisticated principles on how to rank outcomes both in terms of how much moral reason the agent has to desire them and in terms of how much reason, all things considered, she has to desire them. Below, I'll demonstrate how these more sophisticated versions of DRAC can accommodate various intuitive judgments concerning agent-centered constraints.

Let A1-A3 stand for three distinct agents. Let V1-V3 stand for three distinct potential victims. Let  $t_1$ - $t_3$  stand for three distinct and sequential

<sup>44</sup> To consequentialize a nonconsequentialist theory with an absolute constraint against committing murder, we need only replace '500' with ' $\infty$ ' in Mod-SU.

times. And let C1 stand for some set of agent-centered constraints.<sup>45</sup> Assume that whenever I discuss multiple infringements of C1—whether it be multiple infringements by different agents or by the same agent at different times—each infringement of C1 is, in all morally relevant respects, comparable to the others. That is, assume that there are no morally relevant differences between each of the infringements. So, for instance, assume that what each potential victim stands to lose is the same in each instance. And assume that, for each infringement, the agent's moral responsibility for that infringement is the same as it is for all the other infringements and their respective agents.

Now, consider the following schema:

- 4.1** It is impermissible for A1 to infringe upon C1 even if this is the only way that A1 can (short of doing something even worse) ensure that A2 and A3 don't each commit a comparable infringement of C1.

A number of our intuitive judgments are instances of this schema. Here are two examples:

- 4.1a** It is impermissible for A1 to commit murder even if this is the only way that A1 can (short of doing something even worse) ensure that A2 and A3 don't each commit a comparable murder.
- 4.1b** It is impermissible for A1 to break a promise even if this is the only way that A1 can (short of doing something even worse) ensure that A2 and A3 don't each comparably break a promise.

As I've already shown, versions of DRAC, such as Mod-SU, can account for deontic verdicts such as 4.1a and 4.1b. But although consequentialism can account for such deontic verdicts, some philosophers argue that it gives the wrong explanation for them.<sup>46</sup> The consequentialist's

<sup>45</sup> I talk about a set of constraints rather than a single constraint because sometimes a single act-token will violate more than one constraint. Suppose, for instance, that I promised that I wouldn't harm Hanna. In that case, if I murder Hanna by, say, shooting her, then, in shooting her, I violate both the constraint against breaking a promise and the constraint against committing murder.

<sup>46</sup> See, for instance, BROOK 1991 and KAMM 1996, especially pp. 239-247.

explanation for such deontic verdicts lies with, say, the fact that agents ought to have a special concern for their own agency—that is, for what they themselves do as opposed to what they merely allow to happen. Consequently, agents have more moral reason to desire that they refrain from infringing upon a constraint than to desire that each of two others refrain from comparably infringing upon the same constraint.

Contrary to this, Frances Kamm has argued that the explanation for why it is wrong, say, to murder one even so as to prevent two others from committing murder lies with the fact that the one who stands to be murdered for the sake of the two has what Kamm calls a *constraining right* against being murdered—that is, a right not to be murdered even for the sake of minimizing comparable murders by oneself or by others.<sup>47</sup> Of course, the two others also possess this very same right. However, the given agent only comes up against the right of *his* would-be victim, for it is only his would-be victim's right that he is considering infringing. So, on Kamm's view, the rationale for our judgment that it is wrong to murder the one so as to prevent the two others from each committing murder has nothing to do with the thought that the agent ought to have a special concern for his own agency. Rather, Kamm claims that "the agent's own act is special only in that it makes him come up against the constraining right" of his would-be victim.<sup>48</sup> This constraining right acts as a barrier against the permissibility of treating him in certain ways, such as, treating him as a means to the minimization of rights infringements overall.

There are, then, two competing explanations for deontic verdicts such as 4.1a and 4.1b:

- E<sub>1</sub> The explanation lies with the fact that the would-be victim has a certain kind of right—specifically, a constraining right—that acts as a barrier against the permissibility of infringing on that right even for the sake of minimizing comparable infringements of that right by oneself or by others.
- E<sub>2</sub> The explanation lies with some fact other than the one stated in E<sub>1</sub>, such as the fact that agents ought to have a special concern for what they do as opposed to what they merely allow to happen.<sup>49</sup>

<sup>47</sup> See KAMM 1992 and KAMM 1996.

<sup>48</sup> See KAMM 1992, p. 185.

<sup>49</sup> Kamm calls E<sub>1</sub> victim-focused and E<sub>2</sub> agent-focused—see KAMM 1996, especially pp. 239-247. Similarly, Richard Brook calls E<sub>1</sub> patient-based and E<sub>2</sub> agency-based—see BROOK 1991, especially pp. 197-201. Such terminology is misleading, for the proponent of E<sub>1</sub> may be no

Kamm thinks that  $E_1$  offers the better explanation, whereas a consequentialist such as myself must hold that  $E_2$  offers the better explanation. In support of her contention, Kamm cites the following intuitive deontic verdict:

- 4.2** It is impermissible for A1 to kill V3 by introducing V3 to the lethal threat that she previously initiated against V1 and V2 even if this is the only way that A1 can (short of doing something even worse) ensure that V1 and V2 are not killed by this lethal threat.<sup>50</sup>

Imagine, for instance, that A1 earlier planted a bomb that will kill V1 and V2 unless A1 now throws V3 on top of the bomb.<sup>51</sup> If A1 does this, V3 will die, but V1 and V2 will walk away unharmed. If, however, A1 doesn't do this, V1 and V2 will both die, and V3 will walk away unharmed. So A1 must choose either to be the murderer of V3 or the murderer of both V1 and V2. Intuitively, it would be wrong for A1 to throw V3 on top of the bomb, as 4.2 states.  $E_1$  accounts for this so long as we assume that V3 has a constraining right against being murdered. This right constrains A1 from using V3 as a means to minimizing even A1's own infringements of that right.

It may seem that  $E_2$ , by contrast, cannot account for 4.2. After all, if agents ought to have a special concern for their own acts, then shouldn't A1 choose to minimize her own murders by throwing V3 on top of the bomb? The answer is 'not necessarily', as even Kamm recognizes. The proponent of  $E_2$  can account for 4.2 by claiming that what A1 should be most concerned to minimize is not the number of people that she actually murders, but the number of people that she treats as a mere means by attempting to murder them.<sup>52</sup> If this is right, then throwing V3 on top of the bomb is not what A1 should do at all. In doing this, A1 would only be increasing the number of people that she treats as a mere means, for

more focused on the victim (or patient) than the proponent of  $E_2$  is. They may disagree only on what type of right that the victim (or patient) has—that is, they may disagree only as to whether or not the victim's (or patient's) right is of the sort that cannot permissibly be infringed upon even so as to minimize one's own infringements of that right.

<sup>50</sup> KAMM 2007, p. 252.

<sup>51</sup> This is what Kamm calls the "Guilty Agent Case"—see KAMM 1996, p. 242. Scanlon also relies on this case in arguing that  $E_1$  offers the better explanation—see SCANLON 2001, p. 47. And Brook employs a similar sort of case in arguing that  $E_1$  offers the better explanation—see BROOK 1991, p. 197.

<sup>52</sup> See KAMM 1996, p. 244. See also BROOK 1991, pp. 207-208.

nothing that A1 can do now can undo the fact that she has already treated each of V1 and V2 as a mere means. So treating V3 as a means to minimizing the number of people that she murders only adds to the number of people that she treats as a mere means.

To sum up, then, the consequentialist can claim that agents should be more concerned to minimize the number of people that they treat as a mere means by attempting to murder them than to minimize the number of people that they actually murder and that, therefore, A1 ought to prefer the state of affairs in which she lets V1 and V2 die from the lethal threat that she previously initiated to the state of affairs in which she saves those two by treating yet another victim, viz., V3, as a mere means. The consequentialist would, thereby, be employing E<sub>2</sub> to account for 4.2.

In like fashion, the consequentialist can account for the following deontic verdicts, all of which Kamm mistakenly believes will be difficult for the consequentialist to explain:

- 4.3 A1 is morally required to stop A2 from killing V3 even if allowing A2 to kill V3 is the only way that A1 can (short of doing something even worse) ensure that V1 and V2 are not killed by the lethal threat that A1 previously initiated.<sup>53</sup>
- 4.4 A1 is morally required to help V3 escape from A2, who intends to kill V3, even if allowing A2 to kill V3 is the only way that A1 can (short of doing something even worse) ensure that V1 and V2 are not killed by the lethal threat that A1 previously initiated.<sup>54</sup>
- 4.5 A1 is morally required to redirect the lethal threat—e.g., an unstoppable trolley—that is currently threatening both V1 and V2 so that it instead threatens only V3 even if A1 is the one who initiated this lethal threat.<sup>55</sup>

In each case, the consequentialist need only claim that A1 ought to have certain preferences. Thus, to accommodate 4.3, the consequentialist will claim that agents should be more concerned with preventing others from being treated as a mere means than with minimizing the number of deaths that they are responsible for and that, therefore, A1 ought to prefer the outcome in which she stops A2 from killing V3 to the one in which she

<sup>53</sup> KAMM 2007, p. 252.

<sup>54</sup> KAMM 2007, p. 252.

<sup>55</sup> KAMM 1996, p. 244.

prevents V1 and V2 from being killed by the lethal threat that she previously initiated.

To account for 4.4, the consequentialist will claim that agents should be more concerned with helping people escape from would-be murderers than with minimizing the number of deaths that they are responsible for and that, therefore, A1 ought to prefer the outcome in which she helps V3 escape from A2 to the outcome in which she prevents V1 and V2 from being killed by the lethal threat that she previously initiated.

And, to account for 4.5, the consequentialist will claim that, although agents should be more concerned not to introduce any new threats (and not to introduce any new people to existing threats) than to redirect existing threats so that fewer people will die, they should, nevertheless, be concerned to do the latter. Thus, A1 ought to prefer the outcome in which V3 dies as a merely foreseen side-effect of her redirecting the unstoppable trolley onto the fewer number to the outcome in which she allows V1 and V2 to die as the intended effect of her initial act of introducing that threat. Perhaps, the thought would be that although intending to bring about a person's death as a means to achieving one's aims counts as treating that person as a mere means, doing something that has only the foreseeable consequence of bringing about that person's death does not. In any case, it should now be clear that the consequentialist can accommodate intuitions such as 4.1a, 4.1b, and 4.2-4.5.

### **§4.3 Beyond the deontic equivalence thesis: How consequentialist theories can do a better job of accounting for our considered moral convictions than even some nonconsequentialist theories can**

Thus far, I've established that consequentialism can accommodate the same deontic verdicts that Kamm's rights-based deontology can. Kamm might nevertheless object that even if consequentialism can accommodate judgments such as 4.1a and 4.1b, it gives the wrong explanation for them. The consequentialist's explanation is E<sub>2</sub>, whereas Kamm contends that E<sub>1</sub> is the better explanation. Below, I'll argue, contrary to Kamm, that E<sub>2</sub>, not E<sub>1</sub>, is the better explanation. There are two reasons for this. First, E<sub>2</sub> can account for everything that E<sub>1</sub> can account for and more. Second, unlike E<sub>2</sub>, E<sub>1</sub> has counterintuitive implications. I explain each in greater detail below.

Like the proponent of E<sub>1</sub>, the proponent of E<sub>2</sub> can account for deontic verdicts such as 4.1a, 4.1b, and 4.2-4.5. To do so, the proponent of E<sub>2</sub> need only make certain claims about what agents should be most concerned to minimize. To account for 4.2, for instance, the proponent of E<sub>2</sub> need only

claim that agents should be most concerned to minimize the number of people that they treat as a mere means. Unlike the proponent of  $E_1$ , however, the proponent of  $E_2$  is able to account for certain additional deontic verdicts: specifically, instances of 4.1 in which there is no victim associated with  $A_1$ 's infringement of  $C_1$ . To illustrate, consider the *Wrongful Conception Case*. A man named Larry wants very badly to have a biological son, and he can conceive only via coitus, and not through IVF or any other artificial means. But unfortunately for him, he has a rare disease, where only his X-chromosome-carrying sperm can swim. Given this, he can conceive only daughters. But suppose a new pill has just come out on the market that can help Larry get what he wants. The pill will cause a genetic mutation in Larry's sperm, converting all his X-chromosome-carrying sperm into non-swimmers and all his Y-chromosome-carrying sperm into swimmers. An unfortunate side-effect of this mutation, however, is that it causes the sons that he thereby conceives to be born with a terminal genetic disease that induces a sudden and painless death at the age of forty. Larry is so eager to have a biological son that he gladly takes the pill, consequently giving birth to a son, whom he names Matthew. Matthew lives a good life up to the age of forty, which is when he suddenly and painlessly dies of his genetic disease. Had Larry chosen instead to conceive without taking the pill, he would have conceived a daughter, whom he would have named Nicole. Nicole would have lived to the ripe old age of eighty. The quality of her life would have been, on average, no better or worse than Matthew's, but it would have been twice as long.

Let's call the act of taking the pill so as to conceive a boy who will live forty good years instead of a girl who would have lived eighty good years ' $A_T$ ' (the ' $T$ ' standing for 'taking'). Intuitively, it seems that Larry should not, other things being equal, take the pill (thereby performing an instance of  $A_T$ ) even so as to prevent two other men from each taking the pill.<sup>56</sup> Thus, we have, here, the following instance of 4.1:

- 4.1c** It is impermissible for  $A_1$  to perform an instance of  $A_T$  even if this is the only way that  $A_1$  can (short of doing something even worse) ensure that  $A_2$  and  $A_3$  don't each perform a comparable instance of  $A_T$ .

<sup>56</sup> When I conducted an admittedly non-scientific poll, I found that the vast majority of participants who had intuitions such as 4.1a and 4.1b also found 4.1c to be intuitive. For the poll and some analysis, see <http://peasoup.typepad.com/peasoup/2009/02/constraints-agentfocused-or-victimfocused-part-ii.html>.

Unfortunately for the proponent of  $E_1$ , she cannot account for 4.1c. The explanation for 4.1c cannot lie with the fact that the would-be victim has a constraining right, because, in the case of 4.1c, there is no victim, let alone an infringement upon some victim's rights. To see this, consider again the Wrongful Conception Case. And let's assume both that Larry is made somewhat better off as a result of taking the pill and that the only people who are possibly affected by Larry's decision are Mathew, Nicole, and himself. (It's fair to make these assumptions, for they don't seem to affect our confidence in 4.1c.) Given that Larry is somewhat better off having taken the pill, Larry is no victim. Nor is Matthew a victim, for he is not made worse off as a result of Larry's taking the pill. Indeed, had Larry not taken the pill, Matthew would have never existed. And, surely, we should deny that, prior to even being conceived, Matthew had the right not to be conceived just because someone else even happier than himself could have been conceived in his stead. That's a very strange right, and an implausible one at that. So I don't see how Matthew could be the victim or how his rights could have been infringed. This leaves us with Nicole. But Nicole doesn't exist and, thus, has no rights. So she can't be the victim. We can conclude, then, that Larry doesn't infringe on anyone's rights in taking the pill and that, therefore, the explanation for the wrongness of his doing so cannot lie with  $E_1$ .

Of course, just because the proponent of  $E_1$  can't explain something that  $E_2$  can doesn't mean that  $E_2$  is, *all things considered*, the better explanation. But it does mean that, *other things being equal*,  $E_2$  is the better explanation. As a matter of fact, things are not equal, but not in a way that favors  $E_1$ . As it turns out, the problem with  $E_1$  is not just that it cannot explain everything that  $E_2$  can explain. What's worse,  $E_1$  has counterintuitive implications. Consider the following schema:

- 4.6 It is impermissible for A1 to infringe upon C1 once at  $t_1$ , thereby harming V1, even if this is the only way that A1 can (short of doing something even worse) ensure that she doesn't comparably infringe upon C1 either twice at  $t_1$  or once on each of two separate future occasions,  $t_2$  and  $t_3$ , thereby harming V2 and V3.

If  $E_1$  were correct, we should expect 4.6 to be true. The constraining right that V1 possesses would make it impermissible for A1 to treat her as a means to minimizing A1's own infringements of C1. But, as I show below, many people's intuitions in the relevant sort of test case suggests that 4.6 is



false.<sup>57</sup> Of course, we need to be very careful in constructing the relevant sort of test case. It's important that A1's infringing upon C1 once be the *only* way that A1 can (short of doing something even worse) ensure that she doesn't herself *comparably* infringe upon C1 twice. If it's not, then A1 should take the more acceptable means to ensuring that she doesn't comparably infringe upon C1 twice. This means that we should not use as a test case one in which C1 includes an agent-centered restriction, for in order to construct a case in which A1's infringing upon some agent-centered restriction, R1, at  $t_1$  is the *only* way that A1 can ensure that she doesn't *comparably* infringe upon R1 at both  $t_2$  and  $t_3$ , we must imagine that A1 loses control of her actions such that her later infringements of R1 are involuntary.<sup>58</sup> And, in that case, the infringement at  $t_1$  would not be comparable to the infringements at  $t_2$  and  $t_3$ .

Nicolas Sturgeon nicely illustrates the problem:

[I]magine a case in which, unless I deliberately kill one innocent person now, I make it inevitable that I will deliberately kill two tomorrow. How can this be? Perhaps I am supposed to be possessed of a thirst for innocent blood, known to grow steadily in its demands if left unsated, but also to disappear forever once satisfied. Also, although apparently enough in control to make it worth my worrying conscientiously about how to cope with this unwelcome intruder into my psyche, I am somehow debarred from the option of disabling myself—by depriving myself of weapons, turning myself over to the authorities, or whatever. So the case is hard to imagine. (1996, p. 521, note 7)

And it seems even more difficult—perhaps, impossible—to imagine a case in which the only way that I can ensure that I don't *voluntarily* commit two murders in the future is to commit one murder now. If these two future murders are indeed going to be murders that I voluntarily commit, then it must be that I can refrain from committing them—that's part of what it means for them to be voluntary. And if I can refrain from committing these two murders in the future even after refraining from committing murder now, then I should refrain from murder both now and in the future.

Of course, it's easy enough to imagine a case in which if I don't voluntarily commit murder now, I *will* voluntarily commit two future murders. Just imagine that I'm a sane but murderous criminal who

<sup>57</sup> Even if I'm wrong about this, all is not lost for the proponent of E2. She could still account for 4.6 by claiming that agents should be more concerned to refrain from infringing upon any constraints *now* than to refrain from infringing upon more numerous constraints *later*.

<sup>58</sup> I focus here on the diachronic case, because I don't see how it's possible that A1's infringing upon R1 once at  $t_1$  could be the *only* way that A1 can ensure that she doesn't *comparably* infringe upon R1 twice at  $t_1$ .

murders anyone who confronts me, and then imagine that, unless I kill Otto now, Paula and Richard will be undeterred from confronting me in the future. Thus, if I don't kill Otto now, Paula and Richard will confront me in the future, and I will, then, murder them both. But even if that's what I *will* do if I don't murder Otto, this seems like a poor justification for murder. I should not murder Otto so as to prevent myself from murdering Paula and Richard if I could refrain from murder altogether.<sup>59</sup>

To avoid having to imagine that one loses control over one's actions, we should test our intuitions regarding 4.6 using cases in which C1 is a special obligation as opposed to an agent-centered restriction, for where C1 is a special obligation it's possible to imagine a case in which A1 never loses control over her actions and yet it's true that the only way for her to ensure that she doesn't infringe upon C1 twice is for her to infringe upon C1 once. Moreover, when testing our intuitions, we should compare pairs of cases that differ in only one respect: specifically, with respect to whether they're an *inter-agent case* or an *intra-agent case*. *Inter-agent cases* are cases in which the agent can, by infringing upon C1, prevent more numerous others from comparably infringing upon C1. *Intra-agent cases*, by contrast, are cases in which the agent can, by infringing up C1, prevent herself from more extensively infringing upon C1.<sup>60</sup>

Fortunately for us, we don't need to do the testing ourselves, for Lopez et al. (2009) have already done it for us. They have constructed and performed the following experiment:<sup>61</sup>

[H]alf the subjects got the [following] *inter-agent case*.

Three people in Joe's community, Mark, Bill, and Frank, are planning to move over a one week period. Joe has promised Mark that he would help him move. Meanwhile David promised Bill and Frank that he would help each of them move. When the time comes to help, it turns out that the only way for David to keep his promise to Bill and Frank is if Joe drives him, which would require Joe to break his promise to Mark.

They were then asked to indicate agreement (on a 7-point scale) with the following statement:

<sup>59</sup> Actually, things are a bit more complicated than I let on here. In Chapter 7, I explore the issue with greater subtlety—see, in particular, my discussion of actualism, possibilism, and securitism in §§7.1.

<sup>60</sup> I borrow these terms from LOPEZ ET AL. 2009. *Inter-agent cases* are sometimes called *multiple-agent cases*, and *intra-agent cases* are sometimes call *single-agent cases*.

<sup>61</sup> The case that they use, which they call the *Promise-Breaking Case*, is very similar to the one that I presented in PORTMORE 1998, pp. 12-13.

All things considered, Joe should break his promise to Mark, so that David can keep his promises to Bill and Frank.

[The other half of the subjects got the following *intra-agent* case.]

Three people in Joe's community, Mark, Bill, and Frank, are planning to move over a one week period. Joe has promised each of them that he would help them move. When the time comes to do so, Joe realizes that he cannot keep his promises to all three. He can either keep his promise to Mark and break his promises to Bill and Frank, or he can keep his promises to Bill and Frank and break his promise to Mark.

The participants who received this question were asked to rate agreement with the statement:

All things considered, Joe should break his promise to Mark, so that he can keep his promises to Bill and Frank. (LOPEZ ET AL. 2009)

The responses to the inter-agent cases differed significantly from the responses to the intra-agent case.<sup>62</sup> Whereas subjects tended to *agree* with the claim that Joe should break his promise to Mark to ensure that he keeps his promises to Bill and Frank, subjects tended to *disagree* with the claim that Joe should break his promise to Mark to ensure that David keeps his (David's) promises to Bill and Frank. This experiment suggests, then, that whereas people tend to accept 4.1 and think that it is impermissible for an agent to infringe upon a constraint so as to ensure that two others don't each comparably infringe upon that constraint, people tend to reject 4.6 and think that it is permissible to infringe upon a constraint once so as to ensure that one does comparably infringe upon that constraint twice. This gives us good reason to think that E<sub>2</sub>, not E<sub>1</sub>, is the better explanation for deontic verdicts such as 4.1b.<sup>63</sup>

I have shown, then, that not only can the consequentialist accommodate deontic verdicts such as 4.1a and 4.1b but it can do so while

<sup>62</sup> As Lopez et al. (2009) report: "The mean response in the inter-agent version is 2.37 (out of 7); the mean response in the intra-agent version is 4.89. The difference between the conditions is quite significant ( $t(35) = 5.263$ ,  $p < .001$ )."

<sup>63</sup> See LOPEZ ET AL. 2009 for other experiments that support this contention. Particularly interesting is the fact that in one of their experiments subjects tended to agree with the claim that it would be permissible to *lie* to ensure that one keeps two of one's own promises, while they tended to disagree with the claim that it would be permissible to lie to ensure that someone else keeps two of her promises. Thus, it seems that we may infringe not only upon a special obligation but also upon an agent-centered restriction so as to minimize our overall infringements of agent-centered constraints.

providing a more plausible explanation for them than the one that rights-based deontologists such as Kamm provide. Ironically, then, it is the same intuitions (i.e., those that are instances of 4.1) that have led so many philosophers to reject traditional forms of consequentialism that now, given our new understanding of consequentialism as a theory that can accommodate agent-centered constraints, leads us to believe that commonsense morality might best be understood as a consequentialist theory.

#### §4.4 The implications of the deontic equivalence thesis

In the previous section, I argued that consequentialism can account for certain commonsense intuitions that rights-based deontology cannot account for, such as our intuition that Joe should break his promise to Mark to ensure that he keeps his promises to Bill and Frank in the intra-agent version of the Promise-Breaking Case. But if consequentialism and rights-based deontology yield different verdicts in this case, doesn't this mean that the deontic equivalence thesis (DET) is false? The answer is 'no'. Just because certain versions of consequentialism accommodate our *intuitive* verdict in the intra-agent version of the Promise-Breaking Case doesn't mean that other versions of consequentialism cannot accommodate rights-based deontology's *counterintuitive* verdict in the same case. Indeed, other versions can. To do so, the consequentialist need only claim that agents should be more concerned to ensure that they not break any promises *now* than to ensure that they not break more numerous promises *later*. So consequentialism can accommodate rights-based deontology's verdict in this case.

Given this and my arguments in the previous two sections, we should conclude that DET is true. But having done so, we still need to consider what the implications of this thesis are. Of course, one clear implication is that, for any remotely plausible set of deontic verdicts, there is some substantive version of consequentialism that generates it. But some go further and argue that another implication of DET is that every nonconsequentialist theory can be fairly regarded as a form of consequentialism, and that, therefore, we are all consequentialists.<sup>64</sup> In this section, I argue that this is mistaken.

If DET is true, then, for any remotely plausible substantive version of nonconsequentialism, there is some substantive version of

<sup>64</sup> See, for instance, BROWN 2004 and LOUISE 2004.

consequentialism that is deontically equivalent to it. But even if two theories agree as to *which acts are right*, that doesn't mean that they agree on *what makes those acts right*. It's important to note, then, that moral theories are in the business of specifying what *makes* acts right. And so even two moral theories that are extensionally equivalent in their deontic verdicts can constitute distinct moral theories—that is, distinct theories about what *makes* acts right.

Consider Kantianism and its deontically equivalent consequentialist counterpart, which I'll call *Kantsequentialism*.<sup>65</sup> These two theories agree as to which acts are right, but they would seem to disagree on what makes them right. On Kantianism, acts are right in virtue of their being instances of treating humanity as an end in itself. On Kantsequentialism, by contrast, acts are right in virtue of their outcomes not being outranked by that of any available alternative. Thus, they certainly appear to be two different moral theories.

We have to be careful, though, for appearances can deceive. It's clear that these two theories differ in the expressions they use to pick out the right-making features of actions, but this doesn't mean that the two different expressions do in fact refer to distinct right-making features.<sup>66</sup> Consider an analogy. One person might hold that what makes something the same temperature as water boiling at sea level is that its temperature is 100° Celsius, whereas another person might hold that what makes something the same temperature as water boiling at sea level is that its temperature is 212° Fahrenheit. These two people use different expressions to describe the physical property that makes something the same temperature as water boiling at sea level, but they don't disagree on what that property is, for the two expressions refer to the exact same property. Their views are, therefore, mere notational variants on the same view.<sup>67</sup>

Some philosophers suspect that the same is true of nonconsequentialist theories and their deontically equivalent consequentialist counterparts. To be deontically equivalent, theories must be *necessarily* coextensive in their

<sup>65</sup> I borrow this term from Richard Arneson.

<sup>66</sup> I have Jamie Dreier to thank for getting me to see this point. I also thank him for suggesting the following analogy concerning temperature.

<sup>67</sup> As I use the phrase 'mere notational variant', two theories, A and B, about what makes something X are mere notational variants if and only if A and B both hold that what makes something X is the same metaphysical property but differ only in that they use different descriptions (perhaps, synonymous descriptions—perhaps, non-synonymous descriptions) to refer to that same property.

deontic verdicts.<sup>68</sup> Thus, if Kantianism and Kantsequentialism are deontically equivalent, then, necessarily, an act is an instance of treating humanity as an end in itself if and only if its outcome is not outranked by that of some available alternative. But given that the two expressions on each side of the 'if-and-only-if' are necessarily coextensive, we might suspect that they refer to the very same property.<sup>69</sup> And thus we might suspect that Kantianism and Kantsequentialism are mere notational variants on the same theory. Furthermore, if, as DET implies, every nonconsequentialist theory is necessarily coextensive with its deontically equivalent consequentialist counterpart, we might even suspect that *all* nonconsequentialist theories are mere notational variants on their deontically equivalent consequentialist counterparts. And, in that case, we really would all be consequentialists.<sup>70</sup>

But even if we might suspect that many nonconsequentialist theories are mere notational variants on their deontically equivalent consequentialist counterparts, it's clear that this is not true of all nonconsequentialist theories. After all, two theories are not mere notational variants on the same theory if each is true only if the other is false. And some nonconsequentialist theories are true only if their deontically equivalent consequentialist counterparts are false. This is because some nonconsequentialist theories are committed to a certain substantive view about the ranking of outcomes that is incompatible with the one that the consequentialist must adopt in order to yield the same deontic verdicts that that nonconsequentialist theory yields. In such instances, the nonconsequentialist theory will be true only if its deontically equivalent consequentialist counterpart is false. And, in that case, the two can't be mere notational variants on the same view.

To illustrate, consider rule-utilitarianism, and recall that rule-utilitarianism is a nonconsequentialist theory given my convention of using 'nonconsequentialism' as shorthand for '*non-act-consequentialism*'.<sup>71</sup>

<sup>68</sup> It's not enough that the two theories yield the same deontic verdicts about every act that has so far been performed in the actual world. To be deontically equivalent, the two theories must yield the same deontic verdicts in every possible world.

<sup>69</sup> Of course, more would need to be said to confirm this suspicion.

<sup>70</sup> But, as I've shown in another work, if we are all consequentialists, we're also all Kantians, all contractualists, and all virtue theorists, and all at the same time. See PORTMORE 2007, pp. 59-60. However, I no longer think that this is necessarily absurd. It wouldn't be absurd if these are all notational variants on the same view.

<sup>71</sup> If the reader doesn't care for my stipulative definitions and would like an illustration involving a more stereotypically nonconsequentialist theory, then imagine welfarist Rossian pluralism. This is just Ross's deontological view combined with welfarism such that Ross's

In order to consequentialize certain versions of rule-utilitarianism (viz., those that incorporate agent-centered restrictions), the consequentializer must adopt a substantive view about how outcomes rank that is incompatible with the rule-utilitarian's own substantive view about how outcomes rank. Rule-utilitarians are committed to *welfarism*, the view that one outcome is outranked by another if and only if it contains more aggregate utility. Given this commitment to welfarism, the rule-utilitarian must deny that whether one outcome outranks another depends either on who is doing the ranking or on whether the person doing the ranking is a murderer in one outcome but not in the other. But, in order to consequentialize rule-utilitarianism (which, in its most plausible forms, includes an agent-centered restriction against committing murder but not one against allowing murder), this is what its deontically equivalent consequentialist counterpart theory must hold. That is, the deontically equivalent consequentialist counterpart theory must hold that, other things being equal, an agent ought to prefer an outcome in which she refrains from committing murder to an outcome in which she commits murder to prevent five others from committing murder, whereas the rule-utilitarian must hold that the latter is preferable given both that she is committed to welfarism and that welfarism commits her to the view that agents ought always to prefer the outcome with the most utility. So rule-utilitarianism and its deontically equivalent consequentialist counterpart are committed to incompatible views about how outcomes rank, and thus each can be true only if the other is false. But if each can be true only if the other is false, then they can't be mere notational variants on the same view. So we're not all consequentialists. Some rule-utilitarians, for instance, are clearly not consequentialists (that is, they are clearly not act-consequentialists).<sup>72</sup>

#### §4.5 An objection

I'll end this chapter by considering a common objection. Since one of the consequentializer's main aims is to construct a substantive version of consequentialism that comports with our moral convictions, the consequentializer will need to appeal to our moral convictions in determining how outcomes rank. Otherwise, the consequentializer will be unable to ensure that outcomes rank in such a way that when this ranking is combined with consequentialism's criterion of rightness, the result is a

prima facie duty of beneficence is construed as a prima facie duty to maximize aggregate utility.

<sup>72</sup> For a contrary position, see DREIER 2011.

set of deontic verdicts that comports with our moral convictions. But we might wonder whether this is legitimate. That is, we might wonder whether it is legitimate for the consequentializer to adopt one ranking over another on the basis that the former but not the latter yields intuitive deontic verdicts when combined with the following criterion of rightness:

**CR** An act is morally permissible if and only if its outcome is not outranked by that of any available alternative.<sup>73</sup>

In thinking about this question, we should note that there are at least three possible procedures that the consequentializer might employ in determining how to rank the outcomes that are available to a given agent while holding CR constant:<sup>74</sup>

*The Footian Procedure:* Independently of any judgments that one has about how these outcomes rank, come to some considered judgments about the permissibility of the various acts that produce them, and rank their outcomes so that the only outcomes that are not outranked are those that are produced by acts that one judges to be permissible.<sup>75</sup>

*The Anti-Footian Procedure:* Independently of any judgments that one has about the permissibility of the various act alternatives available to the agent, come to some considered judgments about how their respective outcomes rank, and rank the outcomes according to these judgments alone. To ensure that CR holds, maintain that the only

<sup>73</sup> For the purposes of illustration only, I choose maximizing act-consequentialism's criterion of rightness.

<sup>74</sup> Why hold CR constant? The idea is that the consequentialist is committed to CR. So the consequentialist must rank outcomes in a way that's compatible with both CR and her deontic judgments.

<sup>75</sup> Of course, no one has judgments about how outcomes rank in the abstract. But, in this case, the judgment that one outcome outranks another is just the judgment that the agent ought (i.e., has most reason) to prefer the one to the other. We do have such judgments. I think that many would, for instance, judge both that, if my child has just fallen out of the boat, I ought to prefer that she doesn't drown and that, if Odette can rescue only one of two groups of people drowning, then I ought to prefer that Odette rescues the group that includes my child even if the other group contains more people.

I call this the Footian Procedure, because it holds, like Foot, that the evaluative status of an outcome is determined by the deontic status of the act that produces it. More specifically, Foot (1985) held that an outcome is best only if the act that produces it is morally obligatory.



permissible acts are those that produce outcomes that one judges not to be outranked.<sup>76</sup>

*The Coherentist Procedure:* While keeping CR constant, revise one's pre-theoretical judgments both about how these outcomes rank and about the permissibility of the acts that produce them in light of each other (and in light of one's various background beliefs) until wide reflective equilibrium is reached, and then rank the outcomes accordingly.

The Footian Procedure would yield very little, if any, normative advice beyond what our pre-theoretical deontic judgments already tell us. Thus the consequentializer should eschew the Footian Procedure if she wants her resulting substantive theory to provide such additional normative guidance.<sup>77</sup> And, according to some, if a moral theory fails to provide such guidance, then this is a significant failure, as one of the things that we supposedly hope to gain from moral theorizing is the ability to better deal with moral questions about which we disagree or about which we have no confident pre-theoretical judgments.<sup>78</sup> But just because the consequentializer should perhaps eschew the Footian Procedure doesn't mean that she must eschew any appeal to our deontic judgments and adopt the Anti-Footian Procedure instead. The Coherentist Procedure is a

<sup>76</sup> The product of this procedure is what Scanlon (2001) calls "Foundational Consequentialism."

<sup>77</sup> See VAN ROOJEN 2004, pp. 164-169.

<sup>78</sup> See HOOKER 2000, p. 4. Although I agree that it would be *nice* if the correct moral theory was able to help us deal with moral questions about which we disagree or about which we have no confident pre-theoretical judgments, I'm not convinced that the correct moral theory *must* be able to do this. For instance, it seems to me that the divine command theory could be right even if it were unable to provide us with any normative guidance. It might be, for instance, that the only way we have of determining what God's commands are is to appeal to our moral convictions and the fact that God commands us to act in a certain way if and only if our acting in that way is morally obligatory. In that case, the divine command theory could not help us deal with moral questions about which we disagree or about which we have no confident pre-theoretical judgments. But, for all that, it could be the correct moral theory, accurately identifying what it is that makes acts right and wrong. So although we may *hope* to gain from moral theorizing the ability to better deal with moral questions about which we disagree or about which we have no confident pre-theoretical judgments, the only thing that we *must* gain from moral theorizing is a better understanding of what makes acts right and wrong. In any case, the point that I make here is that even if philosophers such as Hooker and van Roojen are correct in thinking that moral theories must provide normative guidance beyond that which our pre-theoretical judgments provide, this doesn't pose a problem for the consequentializing project.

viable alternative that, like the Anti-Footian Procedure, yields normative guidance beyond that which our pre-theoretical deontic intuitions provide.

Unlike the adherent of the Anti-Footian Procedure, the adherent of the Coherentist Procedure is willing to revise her pre-theoretical judgments about how outcomes rank in light of any potentially counter-intuitive deontic verdicts it yields when combined with CR. And, unlike the adherent of the Footian Procedure, the adherent of the Coherentist Procedure is also willing to revise her pre-theoretical judgments about the deontic statuses of actions in light of any potential conflicts that may arise when she combines her pre-theoretical judgments about how outcomes rank with CR. Sometimes the adherent of the Coherentist Procedure will end up revising her deontic judgments in light of her more firmly held evaluative judgments (i.e., her judgments about how outcomes rank in terms of the agent's reasons for desiring them). Other times she will revise her evaluative judgments in light of her more firmly held deontic judgments. The Coherentist Procedure will, therefore, yield both new deontic judgments and new evaluative judgments.<sup>79</sup> Given the viability of the Coherentist Procedure, then, it seems that the consequentializer can appeal to our considered moral convictions when determining how outcomes rank without rendering her view either circular or uninformative.

There is, of course, no guarantee that the Coherentist Procedure will produce a ranking that when combined with CR yields a set of deontic

<sup>79</sup> I'm assuming that sometimes we judge that an act is permissible (or impermissible) even though we judge that its outcome is (or is not) outranked. We might, for instance, judge that although I ought to prefer the outcome in which I hack into the UNOS database and move my sick child to the top of the list for receipt of a donor organ, I ought not do this. At the very least, I'm assuming that sometimes our degree of confidence in our judgment that a given act is impermissible (or permissible) is greater or lesser than our degree of confidence in our judgment that its outcome is (or is not) outranked. If, contrary to these assumptions, our deontic and evaluative judgments are, from the start, in perfect accord with both CR and our background beliefs, then the Coherentist Procedure will not yield any new judgments.

I'm also assuming that we have intuitions about which outcomes should be preferred to which others that are independent of our intuitions about the deontic statuses of the acts that produce them. It seems to me that we clearly do have such independent intuitions, for we have intuitions about which states of affairs are to be preferred to which others even when these states of affairs are not produced by any agent's action. For instance, we have the intuition that were I to learn that some kid at my daughter's elementary school has died in a fire, I ought to prefer that it was not my daughter who died even if, impersonally speaking, it would be better if it were her as opposed to one of her more gifted classmates who died.

verdicts that comports with all, or even most, of our considered moral convictions. But there are reasons to be optimistic. One reason is that it seems that many of our pre-theoretical judgments about how outcomes rank—that is, about how much reason there is to desire these outcomes—are already what’s needed to render consequentialism compatible with our considered moral convictions. To illustrate, consider the following. Imagine that I visit an oracle to have my future told. The oracle’s visions of the future are always accurate even if they are sometimes less than fully specific. The oracle foresees that a year from now my daughter will be drowning in a lake. The only chance I’ll have for rescuing her will be by using the motorboat that’s always docked at the lake’s pier. The oracle also foresees that a year from now Otto’s two daughters will be drowning in a different lake. The only chance he’ll have for rescuing them is by using the motorboat that’s always docked at that lake’s pier. The oracle foresees that only one of us will succeed in our rescue attempt, as he foresees that one of the two motorboats will fail to start, though he can’t tell which one it is. This part of his vision is just too hazy.

Unfortunately, there is no way to prevent what the oracle foretells from coming true; this I have learned from the many who have tried before me. All I can do, then, is hope that I will be the one to succeed. And, intuitively, this seems exactly what I ought to hope for. That is, it seems that I ought to prefer that I will succeed even though it is better, impersonally speaking, if Otto is the one to succeed, since if he succeeds two girls will be saved as opposed to just one. This ranking of outcomes is exactly what’s needed to ground, on consequentialism, a special obligation to save one’s own children even when one could otherwise enable someone else to save more of their own children. If I ought to prefer that I succeed, then, on consequentialism, I should, if given the choice, save my child rather than enable Otto to save two of his children.

Here’s another example. Imagine that I go to the oracle a second time. This time he foretells that in a year’s time I will face the choice either to commit murder so as to prevent two others from committing murder or to refrain from murder and thereby allow two others to commit murder. The oracle foresees that I will be tempted to murder the one to prevent the two others, but he cannot foresee whether I will ultimately give into this temptation. In this case, it seems to me that I ought to hope and prefer that I choose to refrain from murder. And, again, we find that this is exactly what is needed to render consequentialism compatible with our considered moral convictions—in this case, our conviction that there is an agent-centered restriction against the commission of murder. So one reason to be

optimistic about the Coherentist Procedure yielding an evaluative ranking of outcomes that renders consequentialism compatible with our considered moral convictions is the fact that many of our pre-theoretical evaluative judgments are already what they need to be in order to render consequentialism compatible with our considered moral convictions.

Another reason to be optimistic that the Coherentist Procedure will produce a ranking that when combined with CR yields a set of deontic verdicts that comports with most of our considered moral convictions is the fact that our considered moral convictions will in part be shaped by this Coherentist Procedure. That is, sometimes we will on reflection reject our pre-theoretical deontic judgments precisely because they conflict with deontic verdicts that are entailed by CR and our pre-theoretical evaluative judgments. So even if, in some instances, our pre-theoretical evaluative judgments yield deontic verdicts that conflict with our pre-theoretical deontic judgments when combined with CR, we will sometimes on reflection be willing to jettison some of these pre-theoretical deontic judgments in light of the intuitive attractiveness of MR and TCR and their entailment of consequentialism. So, some of the pre-theoretical deontic judgments that might have otherwise made trouble for the consequentializing project will not survive our reflections. And so there is hope that the consequentializing project will succeed in advancing the cause of consequentialism by yielding a substantive version of consequentialism that comports with most, if not all, of our considered moral convictions and without our having to adopt an implausible ranking of outcomes.<sup>80</sup>



In this chapter, I've argued that any plausible nonconsequentialist theory can be consequentialized. This doesn't mean that we are all consequentialists, but it does mean that it is possible to construct a substantive version of consequentialism that comports with our considered moral convictions and, thus, with commonsense morality. Of course, just because it's possible to construct a substantive version of consequentialism that comports with our considered moral convictions doesn't mean that the resulting version of consequentialism is plausible. A given version of

<sup>80</sup> I'm assuming that any ranking that is in wide reflective equilibrium with both our deontic judgments and our background beliefs, including our views about practical reasons and their relationship to morality—I'm thinking of the teleological conception of reasons and moral rationalism, respectively—would not be an implausible ranking.

consequentialism is only as plausible at its underlying ranking of outcomes. But I've argued that there are good reasons to think that the most plausible ranking of outcomes will be the one that yields a substantive version of consequentialism that comports with most, if not all, of our considered moral convictions.

Of course, consequentialism's ability to comport with our considered moral convictions (including our convictions in certain agent-centered options) is predicated not only on the plausibility of a certain ranking of outcomes, but also on the plausibility of adopting DRAC's dual-ranking structure. And DRAC's dual-ranking structure can seem unmotivated. But things are not as they seem. There are, as I'll show in the next chapter, good reasons to adopt DRAC's dual-ranking structure. Thus, in the next chapter, I will continue to defend the plausibility of consequentializing commonsense morality by defending DRAC against this and other potential objections.

## CHAPTER 5

---

# Dual-Ranking Act-Consequentialism<sup>1</sup>

## *Reasons, Morality, and Overridingness*

- 5.1 Some quick clarifications
- 5.2 Moral reasons, overridingness, and agent-centered options
- 5.3 Moral reasons, overridingness, and supererogation
- 5.4 A meta-criterion of rightness and how it leads us to adopt dual-ranking act-consequentialism
- 5.5 Norcross's objection
- 5.6 Splawn's objection
- 5.7 Violations of the transitivity and independence axioms

RECALL THAT, ACCORDING to what I call *dual-ranking act-consequentialism*,

**DRAC** S's performing  $x$  is morally permissible if and only if, and because, there is no available act alternative that would produce an outcome that S has both more *moral reason* and more reason, all things considered, to want to obtain than to want  $x$ 's outcome to obtain.

Again, I call this *dual-ranking act-consequentialism*, because, on this version of consequentialism, the permissibility of an action is a function of how its outcome ranks relative to those of its alternatives on a ranking of outcomes (i.e., the principal ranking) that is in turn a function of two auxiliary rankings: one being a ranking in terms of how much *moral* reason the agent has to want each of the outcomes to obtain, and the other being a ranking in terms of how much reason, *all things considered*, the agent has to want each of them to obtain. Clearly, these two auxiliary rankings are evaluative: that is, they are each a ranking of outcomes in terms of the agent's reasons (or some subset of her reasons—e.g., her moral reasons) for preferring each to the others. Curiously, though, DRAC's principal ranking does *not* seem to be a ranking of outcomes along any evaluative dimension: that is, it does not rank outcomes in terms of how much moral reason, non-moral reason,

<sup>1</sup> Much of the material in this chapter is drawn from PORTMORE 2008a and PORTMORE 2008b.

or all-things-considered reason the agent has to want each of the outcomes to obtain. What, then, motivates adopting this principal ranking and its dual-ranking structure?

It is, of course, quite common for consequentialists to adopt more sophisticated ranking principles in the hopes of reconciling consequentialism with our commonsense moral convictions. Some consequentialists, for instance, claim that the ranking of outcomes is a function of not only the amount of utility that they contain, but also the fairness of the distribution of the utility that they contain. And, by appealing to such distribution-sensitive ranking principles, consequentialism is made consistent with some of our convictions about distributive justice. But what typically makes this more than just an *ad hoc* move is that the more sophisticated ranking principles are motivated by certain axiological intuitions: i.e., that we ought, other things being equal, to prefer that utility be distributed fairly. But what axiological intuitions could possibly motivate the move from traditional act-consequentialism to DRAC—specifically, the move from a single-ranking structure to a dual-ranking structure? Speaking of the principal rankings that one version of DRAC generates for different agents, Ted Sider admits that they do not “correspond to any independently important axiological facts.” He says, “All I claim for their importance is their role in determining moral normative status” (1993, p. 128). In that case, though, the move to a dual-ranking structure can seem entirely unmotivated.

I concede that the move from a single-ranking structure to a dual-ranking structure isn’t motivated by any axiological intuitions—that is, it isn’t motivated by our intuitions about which outcomes agents ought to prefer. Nevertheless, I will argue that the move to adopt a dual-ranking structure *is* motivated by more than just a concern to avoid counterintuitive implications. I’ll start by showing that moral reasons are not morally overriding in that non-moral reasons can, and sometimes do, prevent moral reasons, even those with considerable requiring strength, from generating moral requirements. And I’ll show that, given this, the consequentialist has no choice but to adopt DRAC.<sup>2</sup>

<sup>2</sup> Let me admit upfront that part of my argument for moral reasons not being morally overriding is that this is something that we must hold in order to accommodate certain moral judgments. The point, though, is not to show that the case for adopting a dual-ranking structure hangs free from our moral judgments, but to show that the case for adopting a dual-ranking structure goes beyond the fact that doing so allows consequentialism to avoid counterintuitive implications. For part of the case for adopting a dual-ranking structure lies with the fact that it allows consequentialism to cohere well with

In §5.1, I clarify what it means to say that moral reasons are not morally overriding, drawing a distinction between *moral reasons* and *morally relevant reasons*. Then, in §5.2, I argue that those who wish to accommodate certain typical agent-centered options must deny that moral reasons are morally overriding and accept both that the reason agents have to promote their own self-interest is a non-moral reason and that this reason can, and sometimes does, prevent the moral reason that they have to sacrifice their self-interest for the sake of doing more to promote the interests of others from generating a moral requirement. And, in §5.3, I argue that those who wish to accommodate many typical instances of supererogation must do the same. In §5.4, I show how the conclusion argued for in the previous two sections—i.e., the conclusion that non-moral reasons can, and sometimes do, prevent moral reasons from generating moral requirements—forces us to accept a certain view about how moral and non-moral reasons function together to determine an act's deontic status, and I show that this view forces the consequentialist to adopt DRAC. Lastly, in the remaining sections, I rebut four potential objections to DRAC.

### §5.1 Some quick clarifications

To hold that moral reasons are *morally overriding* is, as you'll recall from Chapter 2, to hold the following thesis:

**MO** If S has more moral reason to perform  $x$  than to perform  $y$  ( $y \neq x$ ), then S is not morally permitted to perform  $y$ . And, thus, if S has most moral reason to perform  $x$ , then S is morally required to perform  $x$ .

If this thesis were true, then it would always be morally wrong to refrain from doing what one has most moral reason to do, even if what one has most moral reason to do is only supported by the most trivial of moral reasons and opposed by the weightiest of non-moral reasons. Suppose, for instance, that someone were to offer me a million dollars to give a lecture

a range of our moral and meta-ethical judgments, including that there are both supererogatory acts and agent-centered options, that non-moral reasons can justify acting contrary to what there is most moral reason to do, that it is always morally permissible to do what there is most moral reason to do, and that it is always morally permissible to do what there is most reason to do, all things (and, thus, all reasons—moral and non-moral) considered.



at a time at which I'd promised to meet with a student. Assume, for the purposes of illustration, that the reason that I have to keep this promise is a moral reason, that the only reason I have to break this promise is that doing so will further my self-interest, and that the reason that I have to further my self-interest is a non-moral reason.<sup>3</sup> If moral reasons are morally overriding, then the non-moral reason that I have to take advantage of this unique opportunity to make a quick million is, no matter how strong, powerless to prevent the moral reason that I have to keep my promise from generating a moral requirement to do so, no matter how trivial the promise, and no matter how weak the moral reason stemming from it.

As I said in Chapter 2, I think that MO is false. Those, like myself, who deny MO do so because they accept one or both of two theses that entail its falsity. After explaining what these two theses are, I'm going to argue for one of them (without denying that the other might be true as well) and conclude that MO is false. But before I can explain these two theses, I need to draw a distinction and explain some new terminology.

Reasons seem to play at least two normative roles in determining an act's deontic status: a requiring role and a justifying role.<sup>4</sup> Reasons can morally require performing acts that it would otherwise be morally permissible to refrain from performing (the requiring role), and reasons can morally justify performing acts that it would otherwise be morally impermissible to perform (the justifying role). Given these two normative roles, reasons have two potentially separable dimensions of strength: *moral requiring strength* and *moral justifying strength*. A reason has moral requiring strength to the extent that it can make it morally impermissible to refrain from performing acts that it would otherwise be morally permissible to refrain from performing, and a reason has moral justifying strength to the extent that it can make it morally permissible to perform acts that it would otherwise be morally impermissible to perform.

We can use the following criteria to determine whether one reason,  $R_1$ , has more moral requiring strength and/or more moral justifying strength than another reason,  $R_2$ :<sup>5</sup>

$R_1$  has more *moral requiring strength* than  $R_2$  if and only if:

<sup>3</sup> As noted in Chapter 2, we are to assume that I must spend all of the million dollars on personal luxuries such that only I will benefit.

<sup>4</sup> They may also play an enticing role—see note 7 below.

<sup>5</sup> These are adapted from Josh Gert's criteria for *rational* requiring strength and *rational* justifying strength. See GERT 2003, especially pp. 15-16.

- (i)  $R_1$  would make it morally impermissible to do anything that  $R_2$  would make it morally impermissible to do.
- (ii)  $R_1$  would make it morally impermissible to do some things that  $R_2$  would *not* make it morally impermissible to do.

$R_1$  has more *moral justifying strength* than  $R_2$  if and only if:

- (i)  $R_1$  would make it morally permissible to do anything that  $R_2$  would make it morally permissible to do.
- (ii)  $R_1$  would make it morally permissible to do some things that  $R_2$  would *not* make it morally permissible to do.

To illustrate the former criterion, consider that, on commonsense morality, the reason one has to refrain from killing an innocent person has more moral requiring strength than the reason one has to prevent an innocent person from dying. This is true in virtue of the following two facts: (i) If it would be immoral to do something—e.g., to push the red button rather than the green button—because it would entail failing to prevent an innocent person's death, then it would also be immoral to do that same thing if it would entail killing an innocent person. (ii) Even though it would be morally permissible to let an innocent person die in order to save one's daughter (as where both are drowning and one has enough time to save only one of the two), it would not be morally permissible to kill an innocent person in order to save one's daughter (as where one's daughter needs that person's heart to live).

To illustrate the latter criterion, consider that, on commonsense morality, the reason one has to save three lives has more moral justifying strength than the reason one has to save one life. This is true in virtue of the following two facts: (i) If it would be morally permissible to do something that would otherwise be immoral—e.g., to break a promise to meet with a student—in order to save a life, then it would also be permissible to do the same in order to save three lives. (ii) Even though it would not be morally permissible to fail to save two lives in order to save just one (assuming that everything else is equal and, thus, that there is no reason that favors saving the one as opposed to the two), it would be morally permissible to fail to save two lives in order to save three.<sup>6</sup>

<sup>6</sup> Of course, John Taurek has argued that it might be permissible to save the one instead of the two if one were to flip a coin to decide which group to save—see TAUREK 1977.

Having clarified what is meant by the phrases ‘moral requiring strength’ and ‘moral justifying strength’, I can now state one of the two theses that entail the falsity of MO:

**NMRJ** Non-moral reasons can, and sometimes do, prevent moral reasons, even those with considerable moral requiring strength, from generating moral requirements. That is, non-moral reasons have moral justifying strength.

According to NMRJ, it is sometimes the case that *non-moral reasons justify* (hence, the abbreviation ‘NMRJ’) acting contrary to what there is most moral reason to perform. Thus, if NMRJ is true, MO is false.

NMRJ is not the only thesis that entails the falsity of MO, and so it is possible to deny both NMRJ and MO. To see how, let’s assume that there are two alternatives to choose from: *x* and *y*. And let’s assume that there is most moral reason to perform *x*, but most non-moral reason to perform *y*. Now, if NMRJ is false, then the non-moral reasons that support performing *y* will be powerless to prevent the moral reasons that support performing *x* from generating a moral requirement to do so. So, if one is going to deny both NMRJ and MO, then one must accept either (A) that there are some, admittedly weaker, moral reasons that support performing *y* and that these weaker moral reasons can prevent the stronger moral reasons that support performing *x* from generating a moral requirement to perform *x* or (B) that the following thesis is true:

**NMRS** Some moral reasons have absolutely *no moral requiring strength* (hence, the abbreviation ‘NMRS’).<sup>7</sup>

The first of these—namely, A—is clearly implausible. For one, the idea that the weaker moral reasons that support performing *y* could prevent the stronger moral reasons that support performing *x* from generating a moral requirement to perform *x* is quite implausible. For another, it seems that, in

Nevertheless, I’m claiming only that, on commonsense morality, it would be impermissible to save the one instead of the two absent some special reason for saving the one.

<sup>7</sup> It may be that some moral reasons are mere enticers: reasons that, morally speaking, count in favor of performing some act but that don’t have any moral requiring strength whatsoever. These moral enticers could make doing what they entice us to do supererogatory, but they could never make doing what they entice us to do obligatory. Moral enticers would be the moral analogue of what Jonathan Dancy (2004) calls enticing reasons.

many instances, there will be no moral reasons to perform  $y$ , and  $A$  will, therefore, be a non-starter. It seems, then, that if we want to reject both MO and NMRJ, we should accept  $B$ —i.e., NMRS. By accepting NMRS, we can account for MO's being false. For if NMRS is true, then  $S$  can have most moral reason to perform  $x$  and yet not be morally required to perform  $x$  simply because those moral reasons that support performing  $x$  have no moral requiring strength.

I think that NMRS is probably true, but I won't argue for it here. Instead, I'll argue for NMRJ. This alone is sufficient to show that MO is false. I'll assume that there are various typical sorts of agent-centered options and argue that, in order to accommodate them, we must accept NMRJ and, thus, accept the view that non-moral reasons can, and sometimes do, affect an act's deontic status.

But let me rebut one immediate objection. Some will say that, by definition, a reason is a moral one if and only if it's relevant to determining an act's deontic status, and thus it is conceptually impossible for a non-moral reason (which, on this definition, is a reason that is not relevant to determining an act's deontic status) to be relevant to determining an act's deontic status. Of course, one could define a 'moral reason' in this way, but I will adopt a different definition so as to preserve an important distinction between reasons that are relevant to determining an act's deontic status and reasons that, morally speaking, count for or against performing some action. I will call any reason that is relevant to determining an act's deontic status a *morally relevant reason*, and I will call any reason that, morally speaking, counts for or against performing some action a *moral reason*.<sup>8</sup>

This is an important distinction to make, because it may be that not all morally relevant reasons are moral reasons. It is possible that some reasons can justify performing acts that it would otherwise be morally impermissible to perform without themselves counting in favor of performing them, morally speaking. For instance, we might think that the fact I would personally gain from breaking a promise is not a moral reason to do so in that it does not, *morally speaking*, count in favor of my doing so.<sup>9</sup>

<sup>8</sup> Moral reasons either have some moral requiring strength or, if they don't, they are mere moral enticers. Moral enticers can make doing what they entice us to do supererogatory, but they can't make doing what they entice us to do obligatory. Thus, a moral reason is a reason that, if sufficiently weighty, could make an act either obligatory or supererogatory. A reason that could only justify—that is, a reason that couldn't make an act obligatory or supererogatory but could only make an act permissible—would be a non-moral reason.

<sup>9</sup> Nor is it the case that the fact that I would personally gain by performing the act counts against my performing it, morally speaking.

Nevertheless, the reason that I have to act for personal gain may be a morally relevant reason, for, perhaps, such a non-moral reason could, if sufficiently weighty, justify my breaking a promise. In any case, it would be a mistake to rule out, by definition, the very real possibility that non-moral reasons (i.e., reasons that, morally speaking, count neither for nor against any action) are relevant to determining an act's deontic status.

Having clarified the relevant terminology, I'll now proceed to argue that, if we wish to accommodate many typical agent-centered options, we must accept NMRJ—that is, that non-moral reasons justify acting contrary to what there is most moral reason to do. And, in the subsequent section, I'll also argue that, if we wish to accommodate many typical supererogatory acts, we must accept NMRJ.

## **§5.2 Moral reasons, overridingness, and agent-centered options**

An agent-centered option exists whenever an agent is morally permitted either to act so as to make things better overall but worse for herself (or others) or to act so as to make things better for herself (or others) but worse overall. If we delete the parenthetical remarks, we get a description of an *agent-favoring option*. If, instead, we both delete the parenthetical remarks and replace each instance of the word 'herself' with 'others', we get a description of an *agent-sacrificing option*.

The following is a typical instance of an agent-centered option (more specifically, an agent-favoring option): an agent has a certain sum of money that she can use either to secure a considerable benefit for herself or to secure a far more considerable net benefit for various needy, distant strangers. Suppose, for instance, that she must choose to use the money that she has saved either to place a down payment on a new home or to help various needy, distant strangers by donating it to Oxfam.

In this and many other typical instances of agent-favoring options, the following four claims hold:

- 5.1 The agent has the choice to act either self-interestedly or altruistically—that is, she has the choice either to promote her own self-interest or to sacrifice her self-interest for the sake of doing more to promote the interests of others.
- 5.2 It is morally permissible for her to act self-interestedly.
- 5.3 It is also morally permissible for her to act altruistically.

- 5.4 The reason that she has to act altruistically is, what I'll call, a *sufficient requiring reason*—a reason that has sufficient moral requiring strength to generate, absent countervailing reasons, a moral requirement to perform the act of which it counts in favor.<sup>10</sup>

That claims 5.1-5.3 must hold is, I take it, incontrovertible, as each follows from the definition of an 'agent-favoring option'. Claim 5.4, however, is in need of justification.

To see that the reason the agent has to act altruistically is, in many typical instances, a sufficient requiring reason, consider the following two cases. In the first case, Fiona is accessing her savings account via the Internet and is about to transfer the entire balance to her escrow company so as to place the necessary down payment on her new home. She must do this if she is to purchase her new home, and she can do this simply by clicking on the `TRANSFER` button. However, there is an alternative. By clicking instead on the `DONATE` button, her savings will be transferred, not to her escrow company, but to Oxfam. Let's suppose, then, that by clicking on the `DONATE` button she will be providing various needy, distant strangers in the Third World with some considerable, indeed potentially life-saving, benefit (e.g., a new well with enough potable water for everyone in the village). Those who accept that there is an agent-centered option in such cases believe that, given the tremendous sacrifice involved, Fiona is not morally required to click on the `DONATE` button.<sup>11</sup> But it seems that they should also accept that the fact that her doing so would produce such a considerable benefit for these distant, needy strangers constitutes a reason of considerable moral requiring strength to click on the `DONATE` button. Indeed, but for the costs involved, it seems that this reason would generate a moral requirement to click on the `DONATE` button.

<sup>10</sup> So a requiring reason is any reason that has some moral requiring strength, and a sufficient requiring reason is a requiring reason that has sufficient moral requiring strength to generate a moral requirement in the absence of countervailing reasons. It may be that all requiring reasons are sufficient requiring reasons, but I draw the distinction so as to allow for the possibility that some requiring reasons are such that individually they have insufficient moral requiring strength to generate a moral requirement but, in conjunction, they have enough combined moral requiring strength to generate a moral requirement—at least, absent countervailing considerations. Given this, we should allow that, in the above criteria for moral requiring and/or moral justifying strength,  $R_1$  and  $R_2$  can stand for sets of reasons and not just for individual reasons.

<sup>11</sup> If the reader believes that, given that lives are at stake, Fiona is morally required to click on the `DONATE` button, then imagine a revised version of the case in which what's at stake is just slightly less than what is needed to generate a moral requirement.

To see this, consider the second case, a variant of the first. In this case, Fiona can transfer the money to her escrow company by clicking on *either* the `TRANSFER` button or the `DONATE` button. In this case, the money that would be donated to Oxfam if Fiona clicks on the `DONATE` button is not hers, but that of a very rich man who has agreed to donate an equivalent sum of *his* money if, and only if, Fiona clicks on the `DONATE` button. Again, in either case, Fiona's money will be transferred to her escrow company, ensuring the purchase of her new home. It's just that by clicking on the `DONATE` button as opposed to the `TRANSFER` button, she also ensures that various needy strangers receive a considerable benefit. Assume that these are the only morally relevant facts.

Surely, in this case, Fiona is morally required to click on the `DONATE` button, for there is no good reason why she should not do so. By clicking on the `DONATE` button, she can purchase her new home while also providing an even more considerable net benefit for a number of others, and she can do so at no cost to herself or anyone else, for assume that if Fiona doesn't click on the `DONATE` button, the sum of money that the rich man would have otherwise donated to Oxfam will instead be burned, benefiting no one. If one thinks that beneficence is only required when the would-be beneficiaries are below a certain threshold of well-being, assume that the various needy, distant strangers that will be helped by Fiona's clicking on the `DONATE` button are below this threshold.

Given that we think that the reason that Fiona has to click on the `DONATE` button gives rise to a moral requirement in the absence of countervailing reasons, we must conclude that it is a sufficient requiring reason. Now, the only relevant difference between this case and the original one is how costly it is for Fiona to help the strangers. But surely it is implausible to suppose that Fiona's reason to help the strangers, or its moral requiring strength, diminishes as the cost of her doing so increases.<sup>12</sup>

<sup>12</sup> See KAGAN 1989, p. 49. Of course, particularists might object that a reason can have a great deal of moral requiring strength in one context (e.g., the context in which the cost of acting altruistically is quite low) but very little to no moral requiring strength in another context (e.g., the context in which the cost of acting altruistically is quite high), for particularists will deny that reasons have any stable valence or strength values across possible contexts. But see GERT 2007a for an interesting and powerful response. In that article, Gert argues that the particularist "cannot merely deny that it makes sense to ascribe *stable* strength values to reasons: values that they keep from context to context. Rather, the particularist must make the blanket claim that talk of the strength of a reason makes no real sense even in a restricted context" (p. 553). This is because, when we assign a strength value to a reason, we are providing a concise representation of the way it affects the deontic statuses of acts across a range of contexts. For instance, if we assign greater moral requiring strength to one reason

Suppose, for instance, that we were to increase gradually the cost of clicking on the `DONATE` button, from no cost at all, to 10 cents, to 20 cents, to 30 cents, and so on. It is not as if there is less and less to be said in favor of Fiona's helping the strangers as the cost of her clicking on the `DONATE` button increases. At least, that is not what the phenomenology of the case tells us, for it feels like a case in which one reason is outweighed by another, not like a case in which one reason is undermined by another. If it were the latter, then once the cost was high enough, Fiona should cease to feel any pull toward clicking on the `DONATE` button. But even when the cost is extremely high, the fact that clicking on the `DONATE` button would help these others continues to count in favor of Fiona's doing so, and with the same force as before. And there's nothing particularly special about this case. So we should conclude that, in many typical instances in which the agent has an agent-centered option, 5.4 is true.

Of course, some moral theorists who endorse agent-centered options could object that although they are committed to certain moral *principles*, they are not committed to there being any moral *reasons*, let alone to there being a sufficient requiring reason to act altruistically in such cases. They might, therefore, deny that they're committed to 5.4. But their moral principles commit them to certain moral reasons, for moral principles entail that certain facts have sufficient requiring strength. For instance, if one accepts the principle of utility (i.e., that agents are required to maximize utility), then one is thereby committed to there being a sufficient requiring reason to promote utility. The principle of utility entails that the fact that some act would promote utility is a sufficient requiring reason to perform that act. Likewise, anyone who accepts a moral principle that entails that agents are obligated to provide needy others with a significant benefit whenever their doing so would be costless both to themselves and to others is thereby committed to there being a sufficient requiring reason to act altruistically in such instances.

So, one of the assumptions that I'm making is that some such moral principle is true. If this assumption were false, then Fiona would not, in the second case, be obligated to click on the `DONATE` button. And if that were right, then I would be wrong in saying that those who wish to

than to another, we are committed to a claim about how these two reasons affect the deontic statuses of acts across contexts: specifically, we must claim that the one reason would make it morally impermissible to do anything that the other reason would make it morally impermissible to do. Since this is precisely the sort of claim that particularists must deny, they must deny that talk of the strength of a reason makes any sense at all, even in particular contexts.



accommodate many typical agent-centered options are committed to 5.4. I doubt, though, that my assumption is false. In any case, I'll be assuming in what follows that some such moral principle is true and that, therefore, anyone who wishes to accommodate many typical agent-centered options must accept 5.4.

Given 5.4, we must ask: What prevents the sufficient requiring reason that the agent has to act altruistically from generating a moral requirement to act altruistically? Clearly, it must be the reason that the agent has to act self-interestedly, as this is the only countervailing reason, and we must cite some countervailing reason, since, given 5.4, we are to assume that the moral reason that the agent has to act altruistically would generate a moral requirement to act altruistically absent countervailing reasons. We must also assume that this countervailing reason to act self-interestedly must have at least as much moral justifying strength as the reason the agent has to act altruistically has moral requiring strength—otherwise, it wouldn't be able to prevent the reason that the agent has to act altruistically from generating a requirement to act altruistically. Lastly, we must assume that this reason to act self-interestedly must have less moral requiring strength than moral justifying strength, for, otherwise, we would end up with a moral requirement to act self-interestedly instead of a moral option to act either altruistically or self-interestedly. This is Shelly Kagan's worry. He says,

If, in some particular case, the balance of morally relevant reasons did not favor promoting the overall good [i.e., acting altruistically] but favored instead promoting the agent's own interests [i.e., acting self-interestedly]—then it seems that these reasons would still go on to generate a moral *requirement*. Admittedly, the agent would not be morally required to promote the overall good, but she *would* be morally required to promote her interests. Yet...[w]hat we were looking for was a defense of a moral *option*, according to which the agent would still be morally permitted (although not required) to do the act with the best results overall. (1994, pp. 338-339)

The solution, as a number of philosophers have pointed out, lies with the fact that the morally relevant reasons that favor acting self-interestedly as opposed to altruistically are non-moral reasons.<sup>13</sup> If such non-moral reasons can prevent the moral reason that the agent has to act altruistically from generating a moral requirement, then what we end up with is a moral option rather than a moral requirement to act self-interestedly, for non-

<sup>13</sup> See BRATMAN 1994 and SLOTE 1991.

moral reasons, by definition, lack any moral requiring strength.<sup>14</sup> Kagan overlooks this possible solution to his worry, because he assumes that the only sorts of reasons that could prevent a moral reason from generating a moral requirement are other *moral* reasons. He says, “since we are concerned with what is required by *morality*, the relevant reasons—whether decisive or not—must be moral ones” (1989, p. 66). But Kagan’s inference is unwarranted; we should not just assume that non-moral reasons are irrelevant with regard to what is required by morality.

Fortunately for the defender of agent-centered options, it is quite plausible to suppose that the fact that performing some act would further one’s self-interest is not *itself* a moral reason to perform it and is, thus, not a reason of any moral requiring strength. Moral reasons are, of course, a proper subset of reasons for action. So if reasons are considerations that count for or against performing some action, then *moral* reasons are considerations that, *morally speaking*, count for or against performing some action. But there is nothing, *morally speaking*, that counts in favor of promoting one’s self-interest *per se*.<sup>15</sup> This is not to say that one never has a moral reason to do what will further one’s self-interest; one often does, as when doing one’s moral duty coincides with promoting one’s self-interest. The claim is only that the mere fact that performing some act would further one’s self-interest does not *itself* constitute a moral reason to perform that act, for the mere fact that performing some act would be in one’s self-interest is never by itself sufficient to make an act obligatory, or even supererogatory.

Consider, for instance, that the fact that I would benefit from getting a massage does not, *morally speaking*, count in favor of my getting one. If I had the opportunity to get one for free and chose instead to do something less beneficial for myself, I could rightly be called foolish or imprudent, but not immoral. And the mere fact that performing some act would further one’s self-interest is not only incapable of making such an act morally obligatory, it’s also incapable of making such an act morally supererogatory. Consider that in those instances in which I am morally required to come to someone’s aid, as where there is a child drowning in a shallow pond, ensuring that I benefit myself in the process (by, say,

<sup>14</sup> If a given reason for action did have some moral requiring strength, it would thereby count in favor of performing that act, *morally speaking*, and would, therefore, be a moral reason.

<sup>15</sup> For a more thorough defense of this claim, see PORTMORE 2003, especially section III.

alerting the news media so that I might receive some reward) would not by itself constitute going above and beyond the call of duty.

Of course, some might object that there are duties to the self, and that such duties show that there is a moral reason to promote one's self-interest. But the idea that there are certain duties to the self is compatible with the thought that there is no moral reason *per se* to promote one's self-interest. To illustrate, take the duty to develop one's talents. It seems that this duty derives not from some duty to promote one's self-interest, but from some duty to make use of certain valuable gifts, and this explains why we are not morally obligated to develop every talent that would be of potential benefit to ourselves.

Take, for instance, the ability to walk on one's hands over great distances. This is not the sort of talent that one is morally obligated to develop. Of course, one might benefit from developing such a talent, as where one wishes to make it into *The Guinness Book of World Records*. But even so, one wouldn't be morally required, but only prudentially required, to develop that talent. Consider also that it would have been wrong for Mozart to have wasted his unique musical gifts even if he would have been just as well off (self-interestedly speaking) being a mason instead of a musician. It seems, then, that the wrongness of wasting such great gifts lies with its wastefulness and not with its effects on the individual's self-interest. So we can admit that people are sometimes required to develop their talents (and, thus, that are duties to the self), but we shouldn't infer from this that the fact that some act would promote one's self-interest constitutes a moral reason for performing it. That is, we should admit that there is sometimes a moral reason to do that which will promote one's self-interest, but deny that the reason one has to promote one's self-interest is itself a moral reason.

Of course, there is no denying that, on some moral theories, such as act-utilitarianism, the reason one has to promote one's self-interest (i.e., to promote one's utility) is itself a moral reason. Nevertheless, the point of this section is to show that those moral theorists who endorse typical agent-centered options are committed to the view that the reason one has to promote one's self-interest is a non-moral reason and that such non-moral reasons can prevent moral reasons from generating moral requirements. What utilitarians think is, therefore, beside the point, since they deny the existence of agent-centered options. Unlike utilitarians, moral theorists who endorse agent-centered options typically give some credence to the sorts of intuitions that I've appealed to above. In any case, it seems that such moral theorists must hold that the reason we have to

promote our self-interest is a non-moral reason if they wish to avoid, as Kagan worries, trading one moral requirement for another. We should conclude, then, that those committed to these quite typical instances of agent-centered options are further committed to:

5.5 The reason the agent has to promote her self-interest by performing the self-interested option is a non-moral reason.

But how can all of 5.1-5.5 be true? The only way is if the following is also true:

5.6 Non-moral reasons can, and sometimes do, prevent moral reasons, even those with considerable moral requiring strength, from generating moral requirements. That is, non-moral reasons have moral justifying strength. (NMRJ)<sup>16</sup>

Unless 5.6 is true, there is no way that 5.1-5.5 could all be true. To see this, consider the following indirect proof. If, contrary to 5.6, non-moral reasons have no moral justifying strength, then the self-interested reason the agent has to perform the self-interested option would be powerless to prevent the

<sup>16</sup> One way this might be true is if moral rationalism (MR) is true. Moral rationalism is, you'll recall, the view that an agent can be morally required to perform an act only if she has decisive reason, all things considered, to do so. If this is right, then a non-moral reason to do something other than  $x$  could prevent a moral reason to do  $x$  from generating a moral requirement to do  $x$  by tipping the balance of reasons, all things considered, in favor of doing something other than  $x$ .

Now, I've just claimed that a non-moral reason (e.g., a self-interested reason) to do something other than  $x$  can prevent a moral reason to do  $x$  from generating a moral requirement to do  $x$  by tipping the balance of reasons, all things considered, in favor of doing something other than  $x$ . David Copp, however, has argued that "there is no standpoint that can claim normative priority over all other normative standpoints and render a definitive verdict on the relative significance of moral and self-interested reasons" (1997, p. 87). He argues that, in cases of conflict between different kinds of reasons, there is no fact as to what a person has most reason to do, all things considered. But, as I see it, moral and self-interested reasons are both reasons with *rational* requiring force. It's just that some moral reasons, and no self-interested reasons, have in addition to this some *moral* requiring force as well. So I don't see the problem in claiming that the rational requiring strength of the self-interested reason that one has to do something other than  $x$  could be greater than the rational requiring strength of the moral reason one has to do  $x$  such that one is rationally required to do other than  $x$ —i.e., has decisive reason, all things considered, to do other than  $x$ . See also MCLEOD 2001 for a persuasive reply to Copp. And see CHANG 2004a and CHANG 2004c for how we might put both moral reasons and self-interested reasons together.

moral reason she has to perform the altruistic option from generating a moral requirement, for, according to 5.5, the self-interested reason she has to perform the self-interested option is a non-moral reason. Clearly, if non-moral reasons have no moral justifying strength, then they are powerless to prevent moral reasons from generating moral requirements. And, given 5.4, we must assume that the agent has a moral reason to perform the altruistic option and that it has considerable moral requiring strength, such that it will generate a moral requirement absent countervailing reasons. But the only countervailing reason in this instance is the reason the agent has to perform the self-interested option, and, as we've just established, this non-moral reason is, assuming the falsity of 5.6, incapable of preventing the moral reason she has to perform the altruistic option from generating a moral requirement to do so. Thus, if we deny 5.6, we are forced to accept that the agent is morally required to perform the altruistic option, and that would mean that we would have to deny 5.2—that is, deny that it is morally permissible for the agent to perform the self-interested option.

So in order to accept all of 5.1-5.5, we must accept 5.6. And since we must accept all of 5.1-5.5 if we want to accommodate many typical instances of agent-centered options, it follows that, if we want to accommodate many typical instances of agent-centered options, we must accept 5.6 as well. That is, we must accept the view that non-moral reasons can, and sometimes do, prevent moral reasons, even those with considerable moral requiring strength, from generating moral requirements (i.e., NMRJ)—assuming, of course, that I'm right to think that we should seek to accommodate such typical instances of agent-centered options.

In the next section, I'll approach the same issue from a different angle, showing that, if we want to accommodate many typical instances of supererogation, we must accept NMRJ.

### **§5.3 Moral reasons, overridingness, and supererogation**

Let me begin by stating two necessary conditions for an act's being supererogatory:

S's performing  $x$  is supererogatory only if there exists some available alternative,  $y$ , such that:

- (a) S is morally permitted both to perform *x* and to perform *y*,  
and
- (b) S has more moral reason to perform *x* than to perform *y*.<sup>17</sup>

Some might deny that *b* states a necessary condition for an act's being supererogatory, claiming instead that an act is supererogatory if and only if it is both morally optional (condition *a*) and morally praiseworthy.<sup>18</sup> While I do not (at least, not in this chapter) want to deny (or assert) that a supererogatory act must be morally praiseworthy, I do want to argue that *b* states a necessary condition.

A supererogatory act is one that goes above and beyond the call of duty. Thus, an agent performs a supererogatory act only if she exceeds the minimum that morality demands—only if she does 'more than she has to' (MCNAMARA 1996a, p. 417). To illustrate, suppose that I promised to bring you some chicken-noodle soup, because you are too sick to eat anything else and are too weak to go out and buy some, let alone make some. Now, the minimum that's required of me, given my promise, is that I bring you some chicken-noodle soup. And let's suppose that I could most easily do this by stopping on the way to your house to buy you some store-bought soup. But suppose that I instead go out of my way to make you some of the best chicken-noodle soup that you've ever had, taking the time to make it from scratch. In spending half of a day making you some home-made chicken-noodle soup, I go beyond the call of duty, doing more than I have to.

<sup>17</sup> I leave open the question of whether there are any further necessary conditions, such as: (c) S's performing *x* is more morally praiseworthy than S's performing *y*. My preference, though, is to use the term 'supererogatory' to denote the pre-theoretic notion of going above and beyond the call of duty (i.e., that of doing more than one has to), and an act can go beyond the call of duty without being praiseworthy, as where the agent performs the act with a bad motive—see MCNAMARA 2011. Thus, conditions *a* and *b* are, I believe, jointly sufficient. In letting the term 'supererogatory' refer to the pre-theoretical notion of exceeding the minimum that morality demands, I follow MCNAMARA 1996A.

As noted in the previous chapter, I've been using the term 'supererogatory' when I should probably be using the term 'superperfecterogatory' instead. A superperfecterogatory act is one that goes above and beyond what perfect duty requires. The term 'supererogatory', by contrast, is best reserved for acts that go above and beyond what all duty (perfect and imperfect) requires. For now, though, I'll ignore this distinction and wait until Chapter 7 before taking up these complicated issues.

<sup>18</sup> Someone else might deny that *b* is a necessary condition, claiming instead that a supererogatory act is one that involves a greater self-sacrifice for the sake of others than is required, whether or not there is necessarily any moral reason for agents to make such self-sacrifices—see, for instance, HARDWOOD 1998. For why this is problematic, see §§4.2.2.

Now, what I did was certainly praiseworthy, but this alone isn't enough to make it supererogatory, for a supererogatory act must go beyond the call of duty and an act can be praiseworthy without going beyond the call of duty. To see this consider the following case, which I borrow from Paul McNamara (2011).

*Soldier on Point:* A soldier is on point this evening. It's her turn to guard the camp. As a general policy and by mutual agreement of everyone in the camp, the soldier on point has the choice of holding either a first or a second position, the first being slightly better with respect to protecting the camp and the second being slightly safer for the soldier on point. This evening the enemy launches a massive assault on the camp in an attempt to overrun it. Despite the grave danger involved and the temptation to run and hide, our soldier faithfully holds the second position, losing her life in the process but providing those back at the camp with sufficient time to prepare and launch a successful counter assault.

In holding the second position, our soldier does the least that she can permissibly do—the minimum that's required of her. Doing anything less than holding the second position (such as, running and hiding) would have been impermissible. But even though our soldier did not go beyond the call of duty (as she would have had she held the first position), her actions are praiseworthy. As McNamara (2011) points out, "To stay on point (in either position) in the face of high chance of death for the sake of others, knowing the advantages of running and having a much better chance of survival, is surely praiseworthy."

This suggests that for an act to be supererogatory it is not sufficient that it be only both morally optional and morally praiseworthy and that a necessary condition for an act's being supererogatory is that it goes beyond the call of duty. And given my arguments from §4.2.2 that a supererogatory act must involve going beyond the call of duty in terms of doing more of something that there is more moral reason to do and not merely in terms of doing more of something that there is no more moral reason to do (such as, perspire), we should think that *b* is a necessary condition for an act's being supererogatory.

A further reason to accept *b* as a necessary condition is that doing so nicely accounts for the normative force that supererogatory acts supposedly have. Supererogatory acts are acts that it would be good, morally speaking, to perform. Indeed, the facts that make an act

supererogatory are presumably considerations that, morally speaking, count in favor of performing that act over any of its non-supererogatory alternatives. But if, contrary to *b*, agents do not have more moral reason to perform a supererogatory act than to perform its non-supererogatory alternatives, then it is hard to see why there is supposedly something that, morally speaking, counts in favor of their doing so. So we should accept that both *a* and *b* constitute necessary conditions.<sup>19</sup>

Once we accept these two necessary conditions, however, supererogation can seem paradoxical since condition *b* appears to be in tension with condition *a*, as James Dreier has explained:

Morality, we are inclined to think, is a matter of what reasons one has *from the moral point of view*. When there is a supererogatory act available, it would be better for you to perform it. So surely you have a reason, from the moral point of view, to perform the act. You may have some reason not to perform it, but at least typically you have no reason *from the moral point of view* [that is, no moral reason] to refrain from it (if you do have such reason, then it will ordinarily be outweighed by the reason you have to perform, because by hypothesis it is better to perform). But now it is hard to see how it could be permissible, from the moral point of view, to refrain from doing something that you have an undefeated reason (from that very point of view) to do. Everything from the moral point of view speaks in favor of your...[performing the supererogatory act], and nothing at all speaks against it. [In] what sense is it "all right," "permissible," "not wrong" to fail [to do so]? There seems to be no sense at all. Supererogation, according to this way of seeing things, turns out to be impossible. (2004, p. 148)

More formally, the paradox of supererogation stems from the following sort of argument:

- 5.7 By definition, *S*'s performing *x* is supererogatory only if there exists some available alternative, *y*, such that: (*a*) *S* is morally permitted both to perform *x* and to perform *y*, and (*b*) *S* has more moral reason to perform *x* than to perform *y*.
- 5.8 If *S* has more moral reason to perform *x* than to perform *y* (*y* ≠ *x*), then *S* is not morally permitted to perform *y*. (MO)
- 5.9 So, whenever condition *b* is satisfied, condition *a* will not be satisfied. And, yet, for an act to be supererogatory, both

<sup>19</sup> I'm not alone in thinking that *b* is a necessary condition for an act's being supererogatory—see, for instance, DREIER 2004, HORGAN AND TIMMONS 2010, RAZ 1975, and ZIMMERMAN 1993.



conditions must be satisfied. Thus, supererogation is conceptually impossible. (From 5.7 and 5.8)<sup>20</sup>

To dissolve the apparent paradox, the supererogationist (i.e., the person who holds that supererogation is conceptually possible) must deny 5.8, explaining why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement to so act.<sup>21</sup> As I noted above, if one wants to deny 5.8 (i.e., MO), then one must accept either or both of the following two theses: (NMRJ) it is sometimes the case that *non-moral reasons justifying* acting contrary to what there is most moral reason to do and/or (NMRS) some moral reasons have absolutely *no moral requiring strength*.

There are, then, two possible (and not mutually exclusive) explanations for why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement to so act. If, on the one hand, NMRJ is false, then the only possible explanation for why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement is that it has insufficient moral requiring strength to generate a moral requirement.<sup>22</sup> Call this *the insufficient-moral-requiring-strength explanation*. If, on the other hand, NMRJ is true, then another possible explanation for why the morally

<sup>20</sup> For a similar account of the paradox of supererogation, see HORGAN AND TIMMONS 2010.

<sup>21</sup> Of course, there can be cases where the reason that supports some particular supererogatory act-token is not morally undefeated, as where there is another supererogatory act-token that there is even more moral reason to perform. But to simplify the presentation of things, let's assume that S's mutually exclusive and jointly exhaustive options concerning what portion of her income to donate to charity are: (x) give more than 10%, (y) give exactly 10%, and (z) give less than 10%. Assume that x is supererogatory, that y is permissible but non-supererogatory, and that z is impermissible. Thus, we must explain why the morally undefeated reasons in favor of S's performing x don't generate a moral requirement for S to perform x.

<sup>22</sup> One might rightly point out that even a morally undefeated reason of considerable moral requiring strength will fail to generate a moral requirement when it is opposed by some moral reason of equal or incommensurate moral requiring strength. But this cannot explain why the morally undefeated reason that favors performing the supererogatory act fails to generate a moral requirement in those instances where it defeats whatever moral reasons there are for performing some permissible non-supererogatory alternative. Moreover, the morally undefeated reason that favors performing the supererogatory act must defeat (not just successfully counter) these moral reasons for performing the permissible non-supererogatory alternative if it is to meet condition b. Thus, assuming that non-moral reasons cannot affect an act's deontic status, the explanation for why the morally undefeated reason for performing the supererogatory alternative fails to generate a moral requirement must be that it has insufficient moral requiring strength.

undefeated reason that favors performing the supererogatory act fails to generate a moral requirement is that there is a non-moral reason to perform some non-supererogatory alternative that prevents it from so doing. Call this *the non-moral-reason explanation*.<sup>23</sup>

Philosophers such as James Dreier and Michael J. Zimmerman overlook the non-moral-reason explanation, because they assume that NMRJ is false.<sup>24</sup> According to them, non-moral reasons are irrelevant in determining the deontic status of an act. Hence, they assume that the insufficient-moral-requiring-strength explanation is the only possible explanation.

In an effort to spell out how the insufficient-moral-requiring-strength explanation might go exactly, Dreier speculates that there might be two moral points of view, one being the point of view of justice and the other being the point of view of beneficence. Dreier further speculates that reasons stemming from justice have considerable moral requiring strength, but that reasons stemming from beneficence have no moral requiring strength. According to Dreier, supererogatory acts are more beneficent, but not more just, than their non-supererogatory alternatives.<sup>25</sup> So although agents have better moral reason to perform a supererogatory act than to perform any of its non-supererogatory alternatives, they are not morally required to do so, for the relevant reasons (i.e., reasons of beneficence) have no moral requiring strength. What an agent is morally required to do is a

<sup>23</sup> Both are sufficient to explain the falsity of MO, and the two are not mutually exclusive. Indeed, I think that we probably need accept both NMRJ and NMRS if we're going to account for all intuitive cases of supererogation. On the one hand, we need, as I explain below, to appeal to NMRJ and the non-moral-reason explanation in order to account for cases such as the original version of the case of Fiona, as described above. And, on the other hand, we need to appeal to NMRS and the insufficient-moral-requiring-strength explanation in order to account for certain other cases, such as the case of Olivia as described in HORGAN AND TIMMONS 2010. For other cases in which we might need to appeal to NMRS and the insufficient-moral-requiring-strength explanation, see FERRY 2009 and SUIKKANEN 2009a.

<sup>24</sup> That Dreier rejects NMRJ is clear from the fact that he thinks that an act's deontic status is a function of solely moral reasons—see the first sentence in the above quote as well as what he says on p. 149 of the same article. Zimmerman, by contrast, is less explicit, but he does say that if there being more moral reason to perform the supererogatory alternative is essential to supererogation, then any theory wishing to accommodate supererogation will have to declare that there are two sets of moral reasons, deontic and non-deontic reasons (or what I am calling moral reasons with, and moral reasons without, sufficient moral requiring strength)—see ZIMMERMAN 1993, pp. 375-6.

<sup>25</sup> Similarly, in note 11, Zimmerman (1993) offers an example where he supposes that reasons of fidelity have considerable moral requiring strength but that reasons of beneficence have little to no moral requiring strength.

function of only those reasons that have moral requiring strength (i.e., reasons of justice), and there is simply no better reason of this sort to perform the supererogatory alternative.

The problem with Dreier's suggestion and, more generally, with the suggestion that the insufficient-moral-requiring-strength explanation is the *only* possible explanation is that both suggestions rest on the mistaken assumption that the reasons that make a supererogatory alternative morally superior to its non-supererogatory alternatives are always moral reasons of insufficient moral requiring strength. To the contrary, it seems that in many typical instances of supererogation the moral reason that favors performing the supererogatory alternative over its non-supererogatory alternatives is a sufficient requiring reason.

To illustrate, recall the case from §5.2 where Fiona must choose between acting so as to secure a considerable benefit for herself by transferring the money from her savings account to her escrow account and acting so as to secure a more considerable benefit for various needy, distant strangers by instead transferring those funds to Oxfam. In this case, her forfeiting the chance to buy a new home and instead donating her savings to Oxfam is supererogatory.

Given that Dreier and Zimmerman insist on the insufficient-moral-requiring-strength explanation, they must presume that, in this case, the moral reason that favors Fiona's transferring the money to Oxfam as opposed to her escrow company is not a sufficient requiring reason. If it were, then her donating the money to Oxfam would, on their view, be obligatory, not supererogatory, because, on their view, whatever non-moral reason she has to purchase a new home is powerless to prevent the sufficient requiring reason that she has to donate the money to Oxfam from giving rise to a moral requirement to so act. But, as we have already seen in §5.2, the moral reason that favors her donating the money to Oxfam is a sufficient requiring reason; absent countervailing reasons, Fiona is morally *required* to donate the money to Oxfam. And there are many other similar cases in which the moral reason one has to perform some beneficent and supererogatory act is a sufficient requiring reason. So if we are to account for many typical instances of supererogation, we are going to have to accept the non-moral-reason explanation, thereby accepting, contrary to Dreier and others, that non-moral reasons, and not just moral reasons, are relevant to the determination of an act's deontic status. That is, we must accept NMRJ.

#### §5.4 A meta-criterion of rightness and how it leads us to adopt dual-ranking act-consequentialism

What we've learned from the previous two sections is that if we want to accept many typical instances of supererogatory acts and agent-centered options (and I do), then we must accept that the deontic status of an action is a function of both moral and non-moral reasons. But what might that function be? Let us call a criterion that specifies this function by providing necessary and sufficient conditions for an act's being morally permissible in terms of both moral and non-moral reasons a *meta-criterion of rightness*. Below, I'll defend the following meta-criterion of rightness, and then argue that this meta-criterion forces the consequentialist to adopt some dual-ranking version of consequentialism (e.g., DRAC):

**MP** S's performing  $\varphi$  is *morally permissible* if and only if there is no available alternative,  $\psi$ , that S has both more requiring reason and more reason, all things considered, to perform.

I'll defend this meta-criterion by arguing for the following five claims, from which MP can be derived:

- 5.10 If S has *optimal requiring reason* to perform  $\varphi$ , then it is morally permissible for S to perform  $\varphi$ .<sup>26</sup>
- 5.11 If S has *optimal reason* (all things considered) to perform  $\varphi$ , then it is morally permissible for S to perform  $\varphi$ .<sup>27</sup>
- 5.12 If it is morally permissible for S to perform  $\psi$  and S has no less requiring reason to perform an alternative,  $\varphi$ , as to perform  $\psi$ , then, unless there is a further alternative,  $\chi$ , that S has both more requiring reason and more reason, all things considered, to perform, it is morally permissible for S to perform  $\varphi$ .

<sup>26</sup> S has optimal requiring reason to perform  $\varphi$  if and only if there is no alternative,  $\psi$ , such that S has more requiring reason to  $\psi$  than to  $\varphi$ .

<sup>27</sup> S has optimal reason to  $\varphi$  if and only if there is no alternative,  $\psi$ , such that S has more reason to  $\psi$  than to  $\varphi$ . By contrast, S has most (or optimific) reason to perform  $\varphi$  if and only if there is no alternative,  $\psi$ , such that S has at least as much reason to  $\psi$  as to  $\varphi$ . So whereas S can have most reason to perform only one act, S can have optimal reason to perform multiple acts, as where more than one act is tied for first place in terms how much reason there is to perform them.

- 5.13 If it is morally permissible for S to perform  $\psi$  and S has no less reason, all things considered, to perform an alternative,  $\varphi$ , as to perform  $\psi$ , then, unless there is a further alternative,  $\chi$ , that S has both more requiring reason and more reason, all things considered, to perform, it is morally permissible for S to perform  $\varphi$ .
- 5.14 It is not morally permissible for S to perform  $\varphi$  if there is an alternative,  $\psi$ , that S has both more requiring reason and more reason, all things considered, to perform.<sup>28</sup>

5.10 is hardly controversial, as it follows rather straightforwardly from the definition of a 'requiring reason'. By definition, only requiring reasons are capable of generating moral requirements. So even if an agent had more non-requiring reasons (where such reasons include non-moral reasons as well as moral reasons with no moral requiring strength) to do other than that which she has optimal requiring reason to do, these reasons couldn't generate a moral requirement to do otherwise. And if the agent isn't morally required to do other than that which she has optimal requiring reason to do, then it follows that it is permissible for her to do what she has optimal requiring reason to do.

5.11 is, by contrast, considerably more controversial, but only if we are unwilling to concede that non-moral reasons can justify acting contrary to what one has most moral reason to do. Once we concede this (and I've argued that we must concede this—that is, I've argued for NMRJ), there is no more plausible claim about when non-moral reasons justify performing an act that would otherwise be impermissible. For if the non-moral reasons in favor of performing a given act don't make that act morally permissible when they successfully counter all other reasons (moral and non-moral), then when? Therefore, we should accept 5.11, acknowledging that it is always morally permissible to do what one has optimal reason to do, all

<sup>28</sup> More formally and accurately, the argument has the following form:  $\forall x \forall y ((Py \ \& \ \sim Ryx) \rightarrow (\sim \exists z (Rzx \ \& \ Azx) \rightarrow Px))$ ,  $\forall x \forall y ((Py \ \& \ \sim Ayx) \rightarrow (\sim \exists z (Rzx \ \& \ Azx) \rightarrow Px))$ ,  $\forall x (\exists z (Rzx \ \& \ Azx) \rightarrow \sim Px)$ ,  $\exists x Px \ \therefore \forall x (Px \leftrightarrow \sim \exists z (Rzx \ \& \ Azx))$ , where ' $Px$ ' stands for 'it is morally permissible for S to perform  $x$ ', ' $Ryx$ ' stands for 'S has more requiring reason to perform  $y$  than to perform  $x$ ', and ' $Ayx$ ' stands for 'S has more reason, all things considered, to perform  $y$  than to perform  $x$ '. Strictly speaking, then, 5.10 and 5.11 are superfluous, although they'll help with the explication below. Also, although not explicitly stated above, I'm assuming that there is some morally permissible alternative available to S (i.e.,  $\exists x Px$ ). I thank G. Shyam Nair, Errol Lord, and especially Peter Marchetto for helping me with this argument.

things considered. In any case, 5.11 is entailed by moral rationalism, which I defended in Chapter 2.

If only 5.10 and 5.11 were true, then agents would have rather limited options. At most, agents would have, in any given situation, only two options: (1) the option to do what they have optimal requiring reason to do and (2) the option to do what they have optimal reason to do, all things considered.<sup>29</sup> (In those situations in which agents have optimal requiring reason to perform precisely one act, an act which they also happen to have optimal reason to do, all things considered, they would have no options at all.) To illustrate, suppose that there is a burning building with twelve children trapped inside. Assume that a man named Alex knows this and has, consequently, called the fire department. Nevertheless, it is clear to Alex that they probably won't arrive soon enough and that all the children will very likely perish unless he acts now and rescues them himself. Unfortunately, Alex lacks the proper fire-fighting equipment, and so there is no way for him to rescue the children without suffering severe injuries in the process. Moreover, we'll assume that the more trips that he makes into the burning building the more severe and life-threatening his injuries will be. He can rescue, at most, two children per trip, one under each arm. Given the great costs associated with each trip, we should assume that what he has most reason to do, all things considered, is to refrain from making any trips into the burning building. But we should also assume that what he has most requiring reason to do is to save all twelve children, making the necessary six trips into the burning building. So, if only 5.10 and 5.11 were true, Alex would be required to make either zero or six trips—that is, to rescue either none or all of the children.<sup>30</sup>

But, surely, if it's permissible for him to refrain from entering the building at all (rescuing no children), then it is permissible for him to enter the burning building once and rescue two of the children. To account for this, we should accept 5.12 in addition to 5.10 and 5.11, for 5.12 says that if it is morally permissible for S to perform  $\psi$  and S has no less requiring reason to perform  $\varphi$  than to perform  $\psi$ , then we should presume that S is also permitted to perform  $\varphi$ . For instance, if it is permissible for Alex to save

<sup>29</sup> Strictly speaking, there could be more than two options if there were more than one act that was tied for first place either in terms of what there is optimal requiring reason to do or in terms of what there is optimal reason to do, all things considered.

In PORTMORE 2003, I defended a view according to which act is permissible if and only if it is either what the agent has optimal requiring reason to do or what the agent has optimal reason to do, all things considered. I now reject this view and accept MP in its stead.

<sup>30</sup> This is what Ted Sider's Self/Other Utilitarianism implies. See SIDER 1993.

no children and he has no less (and, indeed, more) requiring reason to save two, then presumably it is also permissible for him to save the two. The only instances in which this presumption of permissibility will prove false is where S has both more requiring reason and more reason, all things considered, to do something else—say,  $\chi$ . In that case, S should choose to perform  $\chi$ , not  $\varphi$ .

To illustrate, let's suppose: (1) that  $\alpha_1$  is the act-set that consists in standing around outside while waiting for the fire fighters to arrive, rescuing none of the children, (2) that  $\alpha_2$  is the act-set that consists in entering the burning building once and rescuing two of the children, and (3) that  $\alpha_3$  is the act-set that consists in entering the burning building once and rescuing a completely unimportant trash bin.<sup>31</sup> Given 5.11, it is permissible for Alex to perform  $\alpha_1$ ; it is, after all, what he has optimal (indeed, most) reason to do, all things considered—or, at least, so we're assuming. Given 5.12, it is also permissible for Alex to perform  $\alpha_2$ , as he has no less (and, indeed, more) requiring reason to perform  $\alpha_2$  than to perform the permissible  $\alpha_1$ , and there is no alternative that he has both more requiring reason and more reason, all things considered, to perform. He has more reason, all things considered, to perform  $\alpha_2$  than to perform any act that involves taking more trips into the burning building, and he has more requiring reason to perform  $\alpha_2$  than to perform any act that involves taking fewer trips into the burning building. Thus, according to 5.12, it is also permissible for Alex to perform  $\alpha_2$ . But it is not permissible for Alex to perform  $\alpha_3$ , for  $\alpha_2$  is an alternative that Alex has both more requiring reason and more reason, all things considered, to perform. It's clear that there is more requiring reason to perform  $\alpha_2$  than to perform  $\alpha_3$ , for it's clear that there is more requiring reason to save two children than there is to save a trash bin. And it's clear that Alex has more reason, all things considered, to perform  $\alpha_2$  than to perform  $\alpha_3$ , because, even if  $\alpha_2$  and  $\alpha_3$  are on a par in terms of non-moral reasons, it's clear that there is more moral reason and thus more reason, all things considered, to perform  $\alpha_2$ . So although there is a presumption in favor of its being permissible to do  $\varphi$  if there is no less requiring reason to do  $\varphi$  than to do some other permissible alternative,  $\psi$ , this presumption doesn't hold when there is yet another alternative,  $\chi$ , that the agent has both more requiring reason and more reason, all things considered, to perform than to perform  $\varphi$ . And this is just what 5.12 says.

<sup>31</sup> This example is inspired by one of Kagan's—see KAGAN 1989, pp. 16 & 240.

We should also, I believe, accept 5.13. If it is morally permissible for S to perform  $\psi$  and S has no less reason, all things considered, to perform  $\varphi$ , then presumably it is also permissible for S to perform  $\varphi$ . To illustrate, consider again the above case, and this time let's add to the above suppositions the following three: (4) that  $\alpha_4$  is the act-set that consists in making six trips into the burning building, rescuing all twelve children, (5) that  $\alpha_5$  is the act-set that consists in making five trips into the burning building, rescuing ten of the children, and (6) that  $\alpha_6$  is the act-set that consists in making five trips into the burning building, rescuing ten of the children while taking some extra time on the fifth trip to rummage through the garbage, during which time the children are inhaling more and more smoke, increasing their risk of developing respiratory problems. Given 5.10, it is permissible for Alex to perform  $\alpha_4$ ; it is, after all, what he has optimal (indeed, most) requiring reason to do. Given 5.13, it's also permissible for Alex to perform  $\alpha_5$ , as he has no less (and, indeed, more) reason, all things considered, to perform  $\alpha_5$  than to perform  $\alpha_4$  (a permissible alternative), and there is no alternative that he has both more requiring reason and more reason, all things considered, to perform. He has more reason, all things considered, to perform  $\alpha_5$  than to perform any alternative act-set that involves making more trips into the burning building and he has more requiring reason to perform  $\alpha_4$  than to perform any alternative act-set that involves making fewer trips into the burning building. Thus, according to 5.13, it is also permissible for Alex to perform  $\alpha_5$ . But it is not permissible for Alex to perform  $\alpha_6$ , for  $\alpha_5$  is an alternative that Alex has both more requiring reason and more reason, all things considered, to perform. There's more requiring reason to perform  $\alpha_5$  than to perform  $\alpha_6$ , because there's a requiring for Alex to minimize the children's chances of developing respiratory problems.<sup>32</sup> And it's clear that Alex has more reason, all things considered, to perform  $\alpha_5$  than to perform  $\alpha_6$ , because not only is there more moral reason to perform  $\alpha_5$ , but also more non-moral reason to perform  $\alpha_5$ —after all, Alex has a non-moral reason to avoid the unpleasantness of rummaging through the garbage. So although there is a presumption in favor its being permissible to do  $\varphi$  if there is no less reason, all things considered, to do  $\varphi$  than to do some other permissible alternative,  $\psi$ , this presumption doesn't hold when there is yet another alternative,  $\chi$ , that the agent has both more requiring reason and

<sup>32</sup> What if there was just as much requiring reason to perform  $\alpha_6$  as there is to perform  $\alpha_5$ ? In that case,  $\alpha_6$  would be stupid and irrational, but not immoral.



more reason, all things considered, to perform than to perform  $\varphi$ . And this is just what 5.13 says.

Lastly, we should accept 5.14, which says that it is not morally permissible for S to perform  $\varphi$  if there is an available alternative,  $\psi$ , that S has both more requiring reason and more reason, all things considered, to perform. First, if there's more requiring reason for S to perform  $\psi$ , then some reason would have to justify S's performing  $\varphi$  instead. But, by hypothesis, there is more reason, all things considered, for S to perform  $\psi$  than for S to perform  $\varphi$ . So unless we think that the mere fact that S has *some* reason to perform  $\varphi$  is enough to justify doing so even though there is both more requiring reason and more reason, all things considered, to perform  $\psi$ , we should accept 5.14.

Given 5.10-5.14, we should accept MP as our meta-criterion of rightness, for this is what's entailed by 5.10-5.14.<sup>33</sup> Of course, coming up with a meta-criterion of rightness is only the first step along the path toward developing a substantive moral theory. The next step would be to come up with both a substantive account of when there is more requiring reason to perform one act than to perform another and a substantive account of when there is more reason, all things considered, to perform one act than to perform another. Only then would the meta-criterion yield substantive deontic verdicts. One way of doing this, the consequentialist's way, is to hold that an agent's reasons for performing various alternative actions are a function of her reasons for desiring their associated outcomes. When we combine such a teleological conception of practical reasons with the above meta-criterion, we get DRAC's *criterion of rightness*:

**DRAC<sub>cr</sub>** S's performing  $x$  is morally permissible if and only if there is no available act alternative that produces an outcome that S has both more moral reason and more reason, all things considered, to want to obtain than to want  $x$ 's outcome to obtain.<sup>34</sup>

<sup>33</sup> More accurately, I should say that MP is what is entailed by 5.12-5.14 plus the assumption that there is some morally permissible alternative available to S—see note 28.

<sup>34</sup> If there are both requiring and non-requiring moral reasons for desiring an outcome, and if what an agent has requiring reason to do is function of what she has requiring reason to want as opposed to what she has moral reason (in general) to want, then I'll need to replace 'moral reason' with 'requiring reason' in my formulation of DRAC here and elsewhere. In Chapter 7, I suggest that we should indeed make this replacement.

So DRAC's dual-ranking structure is not unmotivated at all. Indeed, the act-consequentialist has no choice but to adopt DRAC if she is to accommodate NMRJ. It might be objected, though, that my arguments for NMRJ presuppose that there are agent-centered options of the sort that commonsense morality takes there to be, the sort of which 5.1-5.5 are true. But I don't see why the consequentialist can't help herself to this given that her main nonconsequentialist rivals all presuppose that there are such agent-centered options and so are committed to the truth of 5.1-5.5. So these non-consequentialist rivals can't object that 5.1-5.5 aren't true. And if they are true, then the consequentialist has a principled reason for making the move to a dual-ranking structure.

Of course, even DRAC isn't all that substantive. There's still very little flesh on its bone. To add a bit more flesh, we would have to give both a substantive account of when there is more moral reason to desire one outcome than to desire another and a substantive account of when there is more reason, all things considered, to desire one outcome than to desire another. One way to do this is to adopt what I've called *Schefflerian utilitarianism* and its *criterion of rightness*:

**SU<sub>α</sub>** S's performing  $x$  is morally permissible if and only if there is no available act alternative that would produce both more utility for others (i.e., for those other than S) and more egoistically-adjusted utility than  $x$  would, where egoistically-adjusted utility includes everyone's utility but adjusts the overall total by giving S's utility, say, ten times the weight of anyone else's.

On SU, agents are never required to promote their own wellbeing, not even when they can do so at no cost to others. But, on SU, agents are sometimes required to promote the wellbeing of others, as when they can do so at no cost to themselves. So whereas the reason that agents have to promote the wellbeing of others has considerable moral requiring strength, the reason that agents have to promote their own wellbeing has absolutely no moral requiring strength. Indeed, the moral justifying strength of the reason that agents have to promote their own wellbeing is ten times that of the reason they have to promote, to the same extent, the wellbeing of others.<sup>35</sup> It is the strong justificatory force of this reason that accounts for SU's ability to accommodate a wide range of agent-favoring options. And

<sup>35</sup> Remember that I picked the number 10 only for the purposes of illustration. The actual multiplier might be different.

it is the complete lack of requiring force associated with the reason that agents have to promote their own wellbeing that accounts for SU's ability to accommodate a wide range of agent-sacrificing options.

Despite SU's ability to accommodate such a wide range of options, a number of philosophers have objected to SU (or, at least, to a theory that is relevantly similar to it). In the next three sections, I'll address these objections and show why they all fail. I do this, not because I ultimately endorse SU, but because these objections to SU apply equally to the theory that I endorse—viz., commonsense consequentialism. So unless I can show that these objections against SU fail, my theory will be in trouble as well. Nevertheless, I'll focus on SU, because it is considerably less complicated than my own theory. This will greatly simplify my discussion of the objections.

### §5.5 Norcross's objection

On SU, there is a *pro tanto* duty of beneficence, for there is always more requiring reason to promote the utility of others than to refrain from doing so. Nevertheless, agents are not always required to promote the utility of others, as sometimes the self-sacrifice involved is too great. To illustrate, suppose that S must choose whether or not to perform  $x$ , where  $x$  is some self-sacrificing, beneficent act. And let's suppose that the difference in overall utility for others if S performs  $x$  as opposed to  $\sim x$  is  $n$ . On SU, S is permitted to refrain from performing  $x$  so long as the amount of utility that S would sacrifice in performing  $x$  is greater than or equal to one-tenth of  $n$ . If, however, the amount of utility that S would sacrifice in performing  $x$  is less than one-tenth of  $n$ , S would be required to perform  $x$ . Interestingly, then, although there can, on SU, be more or less moral reason to perform an act, an act can't be more or less wrong. On SU, right and wrong is an all-or-nothing affair.

Alastair Norcross (2006a, 2006b) has argued that consequentialists such as myself should reject such a theory. He cites the following two reasons. First, a theory that held both that there was a duty of beneficence (e.g., a duty to give some percentage of one's income to charity) and that rightness and wrongness was an all-or-nothing affair "would have to say that there was a *threshold*, e.g., at 10 percent, such that if one chose to give 9 percent one would be wrong, whereas if one chose to give 10 percent one would be right. If this distinction is to be interesting, it must say that there is a *big* difference between right and wrong, between giving 9 percent and giving 10 percent, and a small difference between pairs of right actions [e.g.,

giving 10 versus 11 percent], or pairs of wrong actions [e.g., giving 8 versus 9 percent]" (Norcross 2006b, 41). And Norcross argues that the consequentialist should deny that the difference between giving 8 percent and giving 9 percent is any less significant than the difference between giving 9 percent and giving 10 percent. After all, in each case, the difference in terms of how much good would be done is exactly the same—that is, assuming a fixed income. Thus, if the consequentialist had to choose between convincing only one of two people to give an extra one percent, where one is currently giving 9 percent and the other is currently giving 8 percent, she should be indifferent, for, either way, the result will be that the same amount of additional money goes charity.<sup>36</sup>

Second, Norcross claims that another "related reason to reject an all-or-nothing line between right and wrong is that the choice of any point on the scale of possible options as a threshold for rightness will be *arbitrary*" (Norcross 2006b, 41). Thus, Norcross would claim not only that SU's current multiplier of ten is arbitrary (which, of course, it is), but also that any alternative multiplier that we might replace it with would be equally arbitrary. And since we should reject arbitrary moral distinctions, we should, Norcross would argue, reject SU no matter what multiplier it employs.

Neither of these are good reasons for rejecting SU.<sup>37</sup> To see why, consider that 'right' and 'wrong' (i.e., 'permissible' and 'impermissible') might be like 'worth it' and 'not worth it'. For instance, it might be worth it to trade one's services—let's say an hour of babysitting—for \$10, but not worth it to trade those services for \$9. It's not that there is a bigger difference between \$9 and \$10 than there is between \$8 and \$9. But a \$1 increase in the offered payment from \$9 to \$10 can make all the difference as to whether or not it's worth it to trade one's services for the offered payment. This is true even if a \$1 increase in the offered payment from \$8 to \$9 makes no difference as to whether or not it's worth it to trade one's services for the offered payment. Norcross seems to assume that if the distinction between, say, what's worth it and what's not worth it is to be an interesting one, the difference between \$9 and \$10 would have to be greater

<sup>36</sup> As Rob Lawlor (2009b, p. 104) points out, Norcross is conflating two separate issues here: (1) the issue of whether convincing someone to give an extra one percent would make a difference as to whether that someone acts permissibly and (2) the issue of whether we ought to care about the permissibility of someone's actions over and above what good those actions would do.

<sup>37</sup> For a more detailed and thorough discussion of the flaws in Norcross's arguments, see LAWLOR 2009a and LAWLOR 2009b.

than the difference between \$8 and \$9. But although there is no greater difference between these two pairs of possible payments, the difference between the first pair, but not the second, can make all the difference as to whether or not it's worth it to trade one's services for the offered payment.

Furthermore, there is nothing arbitrary about the \$10 threshold; the threshold is set by the point at which it becomes worth it to trade one's services for the offered payment. Likewise, there needn't be anything arbitrary about the 10 percent threshold; the threshold is set by the point at which it would become contrary to reason to give any more to charity. In both cases, then, the threshold lies where an increase in one of the two competing factors tips the balance from not worth it to worth it or from permissible to impermissible. Whether it's worth it to trade an hour of babysitting for a certain amount of money depends on whether one has at least as much reason to want to receive that amount of money as to want to spend that hour doing something else. Likewise, whether it is permissible to refrain from performing some self-sacrificing and beneficent act depends, according to SU, on whether one has at least as much reason, all things considered, to want to avoid such a self-sacrifice as to want to provide those others with that degree of benefit.<sup>38</sup> In both cases, then, there comes a point when the reasons for acting one way justify acting that way as opposed to the other way. There is nothing at all arbitrary about such a point.<sup>39</sup> So Norcross is mistaken in thinking that the consequentialist can't accept some interesting or non-arbitrary threshold for rightness—that threshold lies at the point at which any further self-sacrifice for sake of others would be supererogatory as opposed to obligatory.

## §5.6 Splawn's objection

<sup>38</sup> I'm assuming both that moral reasons (such as the moral reason one has to benefit others) and non-moral reasons (such as the non-moral reason one has to avoid self-sacrifice) are commensurable and that non-moral reasons sometimes outweigh moral reasons such that what one has most reason to do, all things considered, is contrary to what one has most moral reason to do. Furthermore, I'm assuming that when there is more reason, all things considered, to act contrary to what one has most moral reason to do, it is permissible to do so. That is, I'm assuming both NMRJ and MP, both of which I've argued for above. For more on the commensurability of moral and non-moral reasons, see MCLEOD 2001 and CHANG 2004a.

<sup>39</sup> Rob Lawlor makes the same point—see LAWLOR 2009b, p. 106. I originally made this point and the one above in my lectures on scalar consequentialism, which were posted on 4/4/2008 at <http://www.public.asu.edu/~dportmor/Satisficing%20Consequentialism.pdf>.

Consider the case below, which is depicted in Table 5.1, where an agent named Bonnie has the following three mutually exclusive and jointly exhaustive options:  $a_1$ ,  $a_2$ , and  $A_3$ , where  $A_3$  is an act-type—viz., that of Bonnie’s performing some act other than either  $a_1$  or  $a_2$ .

**Table 5.1**

Let  $U_s(x)$  = the utility that accrues to S if S performs  $x$ , let  $U_{-s}(x)$  = the utility that accrues to others if S performs  $x$ , let  $U_{+s}(x) = U_{-s}(x) + [10 \times U_s(x)]$ , and let  $U(x)$  = the overall utility that is produced if S performs  $x$ .

$x$	$U(x)$	$U_s(x)$	$U_{-s}(x)$	$U_{+s}(x)$	deontic status
$a_1$	-40	10	-50	50	merely permissible
$a_2$	90	-10	100	0	supererogatory
$A_3$	44	-5	49	-1	impermissible

On SU,  $A_3$  is impermissible, because  $a_2$  is an alternative that would produce both (1) more utility for others than  $A_3$  would and (2) more egoistically-adjusted utility than  $A_3$  would—that is, both (1)  $U_{-s}(a_2) > U_{-s}(A_3)$  and (2)  $U_{+s}(a_2) > U_{+s}(A_3)$ . Yet  $A_3$  is considerably better for others than  $a_1$  is ( $U_{-s}(A_3) > U_{-s}(a_1)$ ), and  $a_1$  is a permissible alternative. So why, Clay Splawn wonders, is it permissible for S (i.e., Bonnie) to perform  $a_1$  but not  $A_3$ ? Sure, it’s more costly for S to perform  $A_3$  than it is for S to perform  $a_1$ , but Splawn claims that we should accept the following plausible-sounding principle: “it is always permissible for an agent to make a self-sacrifice so long as that sacrifice does not bring about worse consequences for others” (2001, p. 330). Splawn claims that any theory that fails to accommodate this plausible-sounding principle fails to accommodate the self-other asymmetry.<sup>40</sup> He says, “[I]sn’t that *exactly* the suggestion of the self-other asymmetry that sacrificing one’s own good for the benefit of others is permissible? And isn’t that *exactly* what [ $A_3$ ] is: a considerable self-sacrifice on the part of S in order to bring about some good for others?” (2001, p. 330). And, yet, I claimed in Chapter 4 that SU does accommodate the self-other asymmetry. In response, I’ll argue that Splawn’s plausible-sounding principle is, in fact, *implausible* and that, contrary to what Splawn claims, a

<sup>40</sup> The objection was originally raised against Ted Sider’s Self/Other Utilitarianism, but it applies equally to SU. See SIDER 1993.

theory doesn't need to accommodate this principle in order to accommodate the self-other asymmetry.

To see why the principle is implausible, we'll need to get clear on what precisely it says, specifically, what we're supposed to be comparing when we judge that a self-sacrificing act does or does not "bring about worse consequences for others." One suggestion might be that we're comparing what the total wellbeing for others would be were the agent to interfere with ongoing causal processes by performing the self-sacrificing act with what it would be were the agent to refrain from interfering with ongoing causal processes and omit the self-sacrificing act. But I don't think that this is what Splawn has in mind, and, in any case, the principle is clearly implausible on such an interpretation. Consider again the burning building case from above and suppose that Alex takes only one trip into the burning building, rescuing the trash bin instead of the two children. Did he act wrongly? Well, had he not performed this self-sacrificing act, the consequences would not have been worse for others (on this interpretation), for had he done nothing and just waited for the fire fighters to arrive, the two children would have still died. So, on this interpretation of the principle, taking one trip into the burning building and rescuing the trash bin instead of the two children is permissible, which is just absurd.

A more plausible suggestion, therefore, might be that we are, for each available alternative, to compare what the total wellbeing for others would be were the agent to perform that alternative with what it would be were the agent to perform the self-sacrificing act in question, and so long as there is at least one alternative where others would be no worse off, then the self-sacrificing act is deemed permissible. On this interpretation of the principle, it is permissible for Bonnie to perform  $A_3$  (as Splawn thinks it should be), for there is at least one alternative where others would be no worse off: namely,  $a_1$ . Indeed, not only would others be no worse off, they would in fact be better off:  $U_{-S}(A_3) > U_{-S}(a_1)$ . The problem, though, with this interpretation is that it ignores the fact that  $a_2$  is a more attractive alternative. So although it would be permissible for Bonnie to perform  $A_3$  were  $a_1$  the only available alternative, it is arguably impermissible for Bonnie to perform  $A_3$  given that  $a_2$  is an available alternative. It's true that Bonnie does better by others by performing  $A_3$  instead of  $a_1$ , but she can do even better by others by performing  $a_2$ . And what justification does she have for not doing even better by others by performing  $a_2$ ? Given that Bonnie has more reason, all things considered, to perform  $a_2$  than to perform  $A_3$ , there would seem to be no good reason why Bonnie shouldn't perform  $a_2$  instead of  $A_3$ . Thus Bonnie has no excuse for not doing more for

the sake of others by performing  $a_2$  instead of  $A_3$ . Sure,  $a_2$  is more costly, but, by hypothesis, we're supposing that the reason that Bonnie has to avoid this extra cost is insufficient to defeat the reason Bonnie has to do  $a_2$ —namely, that  $a_2$  will do more good for others. So although it is often permissible to make a self-sacrifice so as to benefit others, it is not permissible to do so when one could have benefited others even more without making any greater of a self-sacrifice than one has decisive reason to make.

So when Splawn asks, "What possible justification could be given for saying [ $A_3$ ] is wrong?" (330), the answer is that there is both more requiring reason and more reason, all things considered, to perform  $a_2$  instead. In choosing  $A_3$  over  $a_2$ , Bonnie would be unreasonably selfish. Bonnie would be benefiting herself by 5 utiles at a cost of 51 utiles for others. So even if we grant, as SU does, that Bonnie has ten times more reason to benefit herself than to benefit others, we must still concede that Bonnie has more reason to benefit others by 51 utiles than she has to benefit herself by 5 utiles. And, thus, we should concede that  $A_3$  is wrong.

What's more, we find that, on this interpretation, Splawn's plausible-sounding principle falls victim to the same counterexample that it did on the first interpretation. On this second interpretation, it's also permissible to enter the burning building and save the trash bin instead of the two children, for waiting for the fire department is an available option that is no worse for others than entering the burning building and rescuing the trash bin is. Indeed, the only way to interpret Splawn's principle so that it doesn't get this implausible result is to hold that we are, for each available alternative, to compare what the total wellbeing for others would be were the agent to perform that alternative with what it would be were the agent to perform the self-sacrificing act, and then to hold that if there is even one alternative where others would be worse off, then the self-sacrificing act is *impermissible*. But, on this interpretation,  $A_3$  would, contrary to what Splawn says, be impermissible, for others are worse off on  $A_3$  than on  $a_2$ . Since Splawn insists that the principle implies that  $A_3$  is permissible, this can't be how Splawn wants us to interpret the principle. So we must accept one of the first two interpretations. But, in that case, the principle is quite contrary to commonsense, implying that it is permissible to enter the burning building once and save the trash bin instead of the two children. We should not, then, think that SU is objectionable in that it violates Splawn's principle. Nor should we conclude that SU fails to incorporate the self-other asymmetry in failing to accommodate Splawn's principle.



### §5.7 Violations of the transitivity and independence axioms<sup>41</sup>

SU holds that an act is morally permissible if and only if its outcome is not outranked by that of any available act alternative. And it holds that one outcome outranks another if and only if it contains both more utility for others and more egoistically-adjusted utility than the other does. Interestingly, then, SU's is-not-outranked-by relation (let's call it 'the  $\succeq$  relation') violates the transitivity axiom, according to which, if  $A \succeq B$  and  $B \succeq C$ , then  $A \succeq C$ .<sup>42</sup> To illustrate, consider, again, the scenario depicted in Table 5.1. None of the outcomes in  $O_3$  (i.e., the set of outcomes corresponding to the set of acts in  $A_3$ ) are outranked by  $o_1$  (i.e.,  $a_1$ 's outcome) and  $o_1$  is not outranked by  $o_2$  (i.e.,  $a_2$ 's outcome), but each of the outcomes in  $O_3$  is outranked by  $o_2$ . So SU's is-not-outranked-by relation (i.e., the  $\succeq$  relation) is intransitive. But this is not problematic. We should not expect the  $\succeq$  relation to be a transitive relation. If the  $\succeq$  relation was equivalent to the at-least-as-good-as relation, then we would expect the  $\succeq$  relation to be transitive. But to say that one outcome is not outranked by another on SU's ranking principle is not to say that the one is at least as good as the other, for SU's ranking principle is not a ranking of outcomes in terms of their goodness. So I don't see why we should expect the  $\succeq$  relation to be transitive.

The transitivity axiom isn't the only axiom of expected utility theory that is violated by SU. On SU, the doesn't-have-a-higher-deontic-status-than relation (let's call it 'the  $\supseteq$  relation') violates the independence axiom: If, for any set of available alternatives in which A and B are the only alternatives,  $A \supseteq B$ , then, for any set of available alternatives in which A and B both appear,  $A \supseteq B$ . To illustrate, suppose that  $a_1$  and  $A_3$  were the only available alternatives. In that case,  $a_1$  would not have a higher deontic status than any of the acts in  $A_3$ , for, where  $a_1$  and  $A_3$  are the only available alternatives, both are, on SU, permissible. But where  $a_2$  is also an available alternative,  $a_1$  does have a higher deontic status than the acts in  $A_3$ . Where  $a_2$  is an available alternative, SU entails both that  $a_1$  is permissible and that  $A_3$  is impermissible. And let me just stipulate that, as I'll use the phrase 'higher deontic status', a permissible act always has a higher deontic status than an impermissible one. Thus the  $\supseteq$  relation violates the independence axiom.

<sup>41</sup> I thank Ted Sider for pointing out the need to address the transitivity and independence axioms.

<sup>42</sup> And the same holds, more generally, for DRAC's is-not-outranked-by relation, but I'll focus on SU for the purposes of illustration.

The thought behind the independence axiom is that adding more alternatives shouldn't affect how two alternatives rank relative to each other in terms of their deontic status; the relative ranking between two alternatives is *independent* of what other available alternatives there are. But whereas we should perhaps expect preferences and utility to obey the independence axiom, we should not expect SU (or any other version of DRAC) to do so. To see why, we need only turn again to my response to Splawn's objection. In that response, I gave good reasons to think that how two alternatives rank relative to one another should depend on what other alternatives are available. Performing any of the acts in  $A_3$  is not unreasonably selfish if  $a_1$  is the only available alternative. But performing one of the acts in  $A_3$  is unreasonably selfish if  $a_2$  is an available alternative. So we should not think that the relative ranking between two alternatives is independent of what other available alternatives there are.



In this chapter, I've addressed a number of objections to DRAC, the main objection being that DRAC's dual-ranking structure is unmotivated. In response to this objection, I've argued that an act is morally permissible if and only if there is no available alternative that the agent has both more requiring reason and more reason, all things considered, to perform. And I've argued that, given this, the consequentialist must adopt DRAC's dual-ranking structure. I've also defended DRAC against four objections: Splawn's objection, Norcross's objection, and two objections concerning violations of the axioms of expected utility theory. There remains, however, at least one other important objection. Shelly Kagan has argued that if one accounts for agent-centered options, as the proponent of DRAC does, by asserting that the reason agents are not always morally obligated to promote the overall good is that they sometimes have more reason, all things considered, to promote their own good, then what you get is a rational requirement for them to promote their own good, when what we wanted was both a moral, and a rational, option for them to pursue either their own interests or the greater good. In the next chapter, I will rebut this objection.

## CHAPTER 6

---

# Imperfect Reasons and Rational Options<sup>1</sup>

- 6.1 Kagan's objection: Are we sacrificing rational options to get moral options?
- 6.2 Imperfect reasons and rational options
- 6.3 Securitism
- 6.4 Securitism and the basic belief
- 6.5 Securitism's suppositions and implications

IN THE PREVIOUS chapter, I argued that in order to account for many typical agent-centered options we must hold that the fact that an act would promote the agent's self-interest is a non-moral reason and that, when this non-moral reason successfully counters the moral (requiring) reason that she has to act altruistically, it morally justifies her acting self-interestedly as opposed to altruistically. But it does so without generating a moral requirement to act self-interestedly, for such non-moral reasons are, by definition, without any moral requiring strength. The result, then, is a moral option to act either altruistically or self-interestedly.

In this chapter, I explain and rebut an important objection to this account of agent-centered options.

### §6.1 Kagan's objection: Are we sacrificing rational options to get moral options?

In response to those who have given this sort of account of agent-centered options (e.g., Michael Slote), Shelly Kagan has raised the following important objection:

Slote is arguing, in effect, that whenever the agent has a moral option, then from the rational point of view the reasons the agent has for favoring her own interests outweigh the reasons that support promoting the greater good [that is, support acting altruistically].... But if this is so, then what if anything prevents these

<sup>1</sup> Some of the material in this chapter is drawn from PORTMORE 2010. However, it is important to note that the view that I endorsed in PORTMORE 2010—viz., the Future-Course-of-Action Theory (or 'FCAT')—is not one that I presently endorse. In its place, I have come to accept *securitism*, which is the view that I describe below.

reasons from grounding a rational *requirement* to favor her interests in each such case? (1991, p. 927)<sup>2</sup>

If the answer is that nothing would, then this would seem to be an unacceptable result, for it would mean that in such cases “it would be rationally forbidden—irrational—to choose to do the morally preferable act” (KAGAN 1991, pp. 927-28). This would seem unacceptable for at least two reasons. First, it seems unacceptable to suppose that, whenever there’s a moral option, performing the morally preferable option is always rationally forbidden.<sup>3</sup> Second, it seems unacceptable to suppose that, whenever there’s a moral option, there’s a rational requirement to perform the self-interested option, for we want to account for the intuition that there is both a moral, *and a rational*, option to act either altruistically or self-interestedly.

Nevertheless, Kagan’s objection is, as stated, a bit too quick, for, contrary to what Kagan says, Slote needn’t claim that, from the rational point of view, the non-moral reason the agent has to act self-interestedly *outweighs* the moral reason the agent has to act altruistically. Slote could instead claim that these two opposing reasons exactly balance out. If so, Slote could plausibly claim that the fact that there is just as much overall reason to act self-interestedly as to act altruistically accounts for the lack of a moral requirement to act altruistically. Moreover, since there is just as much reason, all things considered, to perform the one alternative as there is to perform the other, there is a rational option to do either.

As convenient as this might be, it’s quite implausible to suppose that we could account for a rational option whenever there is a moral option by claiming that the relevant reasons always exactly balance. As Kagan rightly points out, such ties would be too rare to account for the wide range of options we take there to be. To illustrate, suppose that I could save some stranger from having to endure great physical pain by sacrificing \$500.<sup>4</sup> If Slote were to suppose that the reasons for and against sacrificing the \$500 exactly balance out and that this is what accounts for both the moral and

<sup>2</sup> Kagan is responding to SLOTE 1991.

<sup>3</sup> Kagan doesn’t deny that performing the morally preferable (and supererogatory) option is sometimes rationally forbidden—see KAGAN 1991, p. 928. What he finds implausible is the idea that all, or even most, supererogatory acts are rationally forbidden.

The account of supererogation that I gave in PORTMORE 2003 did have the implication that all supererogatory acts are rationally forbidden, and both Michael Byron and Betsy Postow rightfully objected to it for this reason. See BYRON 2005 and POSTOW 2005, pp. 245-253.

<sup>4</sup> I borrow this example from KAGAN 1989, pp. 374-375.

rational option to do either, then, if the tie were broken, there would cease to be an option anymore. Suppose, for instance, that the situation were to change such that there was now slightly more reason for or against making the sacrifice, as where either the amount of pain that the stranger faces or the amount that I would need to sacrifice to prevent it has increased slightly. In that case, it seems that we must admit that there is now most reason, all things considered, to perform one of the two alternatives. Yet we think that making such an altruistic self-sacrifice remains rationally optional even if the strength and/or number of the reasons for or against making that self-sacrifice were to increase slightly. And if there's an option both before and after the strength and/or number of the reasons that support just one of the two alternatives increases, we cannot account for such options by supposing that the relevant reasons exactly balance out.

So Kagan's objection amounts to the following. If the reasons in favor of each alternative are not exactly balanced out, then it seems that one or the other alternative will be the one that there is most reason, all things considered, to perform, and, if so, that alternative will be rationally required even if not morally required. And the reasons in favor of each alternative only rarely exactly balance out. Therefore, this strategy for defending agent-centered options involves trading, in most instances, a moral requirement for a rational requirement, when what we were looking for was both a moral, and a rational, option to act either altruistically or self-interestedly.

It's important to realize, though, that the problem of accounting for rational options given the implausibility of supposing that the relevant reasons always exactly balance out is a difficult philosophical problem that faces anyone interested in accounting for what Joseph Raz calls:

*The Basic Belief:* In most typical choice situations, the relevant reasons do not require performing one particular alternative, but instead permit performing any of numerous alternatives.<sup>5</sup>

It seems, for instance, that I could now, in accordance with reason, do any of the following: watch TV, read a novel, practice the piano, volunteer for Oxfam, work on this book, play with my daughter, or prepare for my next

<sup>5</sup> See RAZ 1999, p. 100. Raz calls it the *basic* belief, because it seems to be sufficiently entrenched in our commonsense thinking that we should give it "credence unless it can be shown to be incoherent or inconsistent with some of our rightly entrenched views" (1999, p. 100). As he sees it, then, the basic belief is one of our starting points, and our task is to explain it rather than to defend it. I agree.

lecture. But how are we to account for this? On the face it, it seems that there could be such rational options only if there were exactly equal reason to perform each of the optional alternatives, and yet it is difficult to believe that such is the case.<sup>6</sup> For instance, it is difficult to believe that I have just as much reason to watch TV as to either volunteer for Oxfam or work on this book. After all, volunteering for Oxfam seems vastly superior to watching TV in terms of the amount of impersonal good that it would do, and working on this book seems vastly superior to watching TV in terms of the amount of good it would do me—assume that I'm not at this moment too tired to work productively on this book.<sup>7</sup> Moreover, the fact that, in many of these choice situations, the relevant alternatives remain rationally optional even after there has been an increase in the number and/or strength of the reasons that favor just one of the alternatives shows that their optional status could not have been due, in the first place, to a perfect balance of reasons (GERT 2008, p. 14). For instance, it seems that it would still be rationally permissible for me to continue to work on this book even if there were a slight increase in the strength and/or number of reasons that favor just one of the other options, as where, say, Oxfam institutes a new policy of giving each volunteer a delicious cookie. So the puzzle is to explain how, in most choice situations, there could be so many rationally optional alternatives if, as seems to be the case, there is not exactly equal reason to perform each one of them.

It seems, then, that the problem of accounting for rational options is not any more serious for the defender of agent-centered options than it is for anyone else who accepts the basic belief. And if the problem is not specific to the above account of agent-centered options, then it is no

<sup>6</sup> Here and throughout the book, my focus is on objective rationality, not subjective rationality. An act is objectively irrational if and only if there are decisive reasons for the agent to refrain from performing it—that is, if and only if the agent objectively ought to refrain from performing it (see §1.4). The objective rational status of an act is, then, purely a function of the reasons there are for and against performing it, irrespective of whether or not the agent is aware of them. By contrast, the subjective rational status of an act depends, not on what reasons there are, but on what reasons the agent takes there to be, or, alternatively, on the practical mental functioning of the agent—see PARFIT 2008 and GERT 2004, respectively. I will have more to say about this distinction in §6.3 below. The reader should assume that, unless otherwise qualified, I am talking about objective rationality whenever I use the word 'rational' or any of its variants.

<sup>7</sup> Indeed, it seems that we *know* that watching TV isn't the alternative that I have most reason to perform. Thus, our belief that the relevant reasons permit my watching TV can't be explained in terms of any kind of epistemic uncertainty as to whether or not it's the alternative that I have most reason to perform.

objection to this account that it encounters the same problem that anyone wishing to account for rational options faces. Moreover, since this problem is quite general, those who endorse this account of agent-centered options, such as myself, can appeal to the same potential solutions that others appeal to in accounting for the basic belief, and there are, as we'll see presently, quite a few of them.

There are, in fact, at least four possible solutions to the problem of accounting for the basic belief. First, we could adopt a satisficing conception of rationality, where it is rationally permissible to perform any act that is supported by sufficiently weighty reasons. Given this conception of rationality, we could account for the basic belief by claiming that, in most choice situations, there are sufficiently weighty reasons to perform all of the seemingly rationally optional alternatives. Moreover, the defender of agent-centered options could hold that although the non-moral reason that the agent has to act self-interestedly outweighs the moral reason that she has to act altruistically (thereby preventing it from generating a moral requirement), there is still a rational option to do both, for both are supported by sufficiently weighty reasons.

Second, we could claim that reasons for action have not only two separable dimensions of moral strength but also two separable dimensions of rational strength: *rational requiring strength* and *rational justifying strength*. Roughly speaking, a reason has rational justifying strength to the extent that it can make it rationally permissible to perform acts that it would otherwise be irrational to perform, and a reason has rational requiring strength to the extent that it can make it irrational to refrain from performing acts that it would otherwise be rationally permissible to refrain from performing.<sup>8</sup> If we were, then, to claim that, in most choice situations, some of the relevant reasons have a great deal of rational justifying strength but very little to no rational requiring strength, we could thereby account for the basic belief.

Interestingly, Joshua Gert argues that, in comparison to self-interested reasons that pertain to avoiding non-trivial harms, altruistic reasons have little to no rational requiring strength, but just as much rational justifying strength when the amount of harm at stake for others is comparable to the amount of harm at stake for oneself.<sup>9</sup> Thus, on Gert's view, one is rationally permitted to sacrifice one's own life to save the lives of two others even

<sup>8</sup> See GERT 2003 and GERT 2004.

<sup>9</sup> See GERT 2004.

though one would not be rationally required to sacrifice much, if anything, to save those two lives.

If this is right, then the defender of agent-centered options can claim that, given its considerable *moral* justifying strength, the non-moral reason that the agent has to safeguard her own self-interest prevents the moral reason that she has to act altruistically from generating a moral requirement to act altruistically, but that it does so without thereby generating a rational requirement to act self-interestedly, for its *rational* requiring strength is exceeded by the rational justifying strength of her moral reason to act altruistically. And although the moral reason she has to act altruistically has sufficient rational justifying strength to make acting altruistically rationally permissible, it has insufficient rational requiring strength to make acting altruistically rationally required, for, as Gert claims, altruistic reasons have very little or no rational requiring strength at all. So, on Gert's view, we end up with a rational option to act either altruistically or self-interestedly.

Third, we could claim that, in most choice situations, the relevant reasons are either not at all, or only very roughly, comparable (or commensurable).<sup>10</sup> When two competing reasons are incomparable, they neither defeat one another nor exactly balance out (and, furthermore, they would not be on a par). And since whenever the competing reasons for various act alternatives fail to defeat each other, it accords with reason to perform any one of them, claiming that, in most choice situations, the relevant reasons are incomparable and thus fail to defeat each other, allows us to account for the basic belief. If this right, then the defender of agent-centered options could account for rational options by making the following claims: (1) the sorts of reasons that support supererogatory acts are typically altruistic (or, alternatively, impartial) reasons, (2) the sorts of reasons that support their permissible non-supererogatory alternatives are

<sup>10</sup> Two types of reasons,  $R_1$  and  $R_2$ , are *wholly incomparable* if and only if there is, for *no pair* of token instances of  $R_1$  and  $R_2$ , some truth as to whether the one is better than, worse than, or just as good as (or on a par with) the other. And two types of reasons,  $R_1$  and  $R_2$ , are *only roughly comparable* if and only if there is, for *only a few pairs* of token instances of  $R_1$  and  $R_2$ , some truth as to whether the one is better than, worse than, or just as good as (or on a par with) the other. Just how rough the comparability of the two types of reasons is in proportion to how few pairs there are for which there is some truth as to whether the one is better than, worse than, or just as good as (or on a par with) the other. The notions of 'wholly incomparable' and 'roughly comparable' come from PARFIT 2008. However, the definitions are mine.

I use 'comparable' and 'incomparable' as synonyms for 'commensurable' and 'incommensurable'.



typically self-interested (or partial) reasons, and (3) these two types of reasons are not at all, or only very roughly, comparable. Indeed, Henry Sidgwick (1966) made these three claims. He held that impartial reasons and self-interested reasons are wholly incomparable, and this led him to accept a kind of *dualism of practical reason*, where it is always objectively rational to do either what would be impartially best or what would be self-interestedly best.<sup>11</sup> And, recently, Derek Parfit has argued that some form of dualism about practical reason is correct, but that Sidgwick was wrong to think that impartial and self-interested reasons are wholly incomparable. Parfit argues that they are instead only roughly comparable.<sup>12</sup> In either case, though, there would, in most instances, be a rational option to act either altruistically or self-interestedly.

Fourth, we could account for the basic belief by supposing that the relevant reasons are *imperfect reasons*—reasons that do not support performing any specific alternative, but instead support performing any of the alternatives that would each constitute an equally effective means of achieving the same worthy end. Of course, it may seem that acting self-interestedly and acting altruistically cannot plausibly be interpreted as two ways of pursuing the same worthy end, nor as two ways of acting on the same reason. But things are not as they seem. If we recognize that our agency is extended over time and that, at a more general level of description, we can choose between two temporally extended courses of action, one that includes acting self-interestedly now and acting altruistically later and another that includes acting altruistically now and acting self-interestedly later, then it does seem plausible to construe these two courses of action as two ways of trying to achieve the same thing: specifically, a reasonable balance between altruistic pursuits and self-interested pursuits over time. If we think that any reasonably choice-worthy course of action will contain both altruistic acts and self-interested acts, then it doesn't matter whether I perform an altruistic act now and a self-interested act later or vice versa so long as, in the end, I perform a sufficient number of each.<sup>13</sup>

To sum up, then, Kagan's objection poses no more serious a problem for the defender of agent-centered options than it does for anyone else

<sup>11</sup> See SIDGWICK 1907.

<sup>12</sup> See PARFIT 2008.

<sup>13</sup> 'Reasonably choice-worthy' can be given either a maximizing or a satisficing interpretation. On the maximizing interpretation, a course of action is reasonably choice-worthy if and only if it is maximally choice-worthy. On the satisficing interpretation, a course of action is reasonably choice-worthy if and only if it is sufficiently choice-worthy.

concerned to account for the basic belief. Moreover, there are, as I've shown, a number of potential solutions to the problem of accounting for the basic belief, and the defender of agent-centered options could potentially avail herself of any one of them. Now, I haven't assessed whether any of these proposed solutions is satisfactory. But whether or not a satisfactory solution can be found, the defender of agent-centered options needn't fear Kagan's objection. If, on the one hand, a satisfactory solution can be found, then it seems likely that the defender of agent-centered options will be able to avail herself of that solution and thereby meet Kagan's objection. If, on the other hand, a satisfactory solution is nowhere to be found, then the defender of agent-centered options will have to admit, as Kagan worries, that there will be a rational option to act either altruistically or self-interestedly only in those rare instances in which the reasons supporting each alternative exactly balance out. But, in that case, Kagan's objection cuts no ice, for, if the basic belief is something that we must ultimately abandon, then everyone (whether they endorse agent-centered options or not) will have to admit that there are rational options only when the relevant reasons exactly balance out.

Since, as I've just argued, the defender of agent-centered options needn't be concerned with whether or not a satisfactory solution to the problem of accounting for the basic belief can be found, it may seem that I should just stop here and remain neutral as to whether or not there is such solution and, if so, what it is. But such a tactic, although probably legitimate, is one that many readers would find unsatisfying. After all, some potential solutions to the problem of accounting for the basic belief might cohere better with the claims and arguments in this book than others. So it would be good to know which solution I advocate and to what extent it coheres with the other theses of the book. Thus, I will, below, try to develop and defend what I take to be the best solution to the problem of accounting for the basic belief. In the process, I'll be defending a particular theory of (objective) rationality: viz., *securitism*. Securitism will not only serve to account for the basic belief and for the fact that our moral options are typically rationally optional as well, but it will also serve as a model for my favored theory of morality – viz., *commonsense consequentialism*. As we'll see in the next chapter, securitism and commonsense consequentialism share many of the same structural features. So explaining and defending securitism in this chapter will make it easier for me to explain and defend commonsense consequentialism in the next. Besides, part of the case for commonsense consequentialism is that coheres well with other theses argued for in this book, including securitism.

## §6.2 Imperfect reasons and rational options

Often times, what an agent has most reason to do is determined by facts about what she has most reason to achieve. But such facts often fail to support any specific alternative, for there is often more than one way to achieve the same result. To illustrate, consider the following three examples. First, the fact that I need to get to the airport provides me with a reason to take any of the following equally attractive means to getting there: a bus, a taxi, or a train—assume that these are all equally attractive given their comparative cost, comfort, and convenience.<sup>14</sup> Second, the fact both that I have a paper that needs to be finished in a month's time and that I will most likely finish it by then if and only if I implement a policy of spending two hours a day working on it provides me with a reason to spend two hours (any two hours) each day over the next month working on it. And, third, the fact that I need to spend one of the next two days grading exams and the other painting the backyard fence in order to meet certain important deadlines provides me with a reason either (1) to spend tomorrow grading exams and the next day painting the backyard fence or (2) to spend tomorrow painting the backyard fence and the next day grading exams.

These three sorts of reasons are what I call *imperfect reasons*, for they are analogous to imperfect duties in that they allow for significant leeway in how one chooses to comply with them.<sup>15</sup> Just as the imperfect duty of beneficence requires only that one be beneficent to a certain extent but leaves it up to one's discretion to whom and on which occasions to be beneficent, imperfect reasons speak in favor of achieving some worthy end but leave it up to one's discretion which of the equally attractive means to achieving this end to take.

Imperfect reasons arise in what Joshua Gert calls *multiple-option cases*, cases in which there is more than one equally attractive means to achieving the same worthy end (2003, p. 10). In multiple-option cases, it is rationally permissible to take any of the equally attractive means to achieving this end, for the relevant reasons are imperfect reasons, which do not support performing any particular act, but instead support performing any of the

<sup>14</sup> I borrow this example from GERT 2003, p. 10.

<sup>15</sup> Jonathan Dancy calls such reasons "unfocused reasons," and Raz calls a subclass of such reasons "reasons that are not time-specific"—see Dancy (2004, p. 100) and Raz (1999, p. 100). I first introduced the notion of an imperfect reason as well as the idea that it could be employed to account for rational options in PORTMORE 2000.

acts that would each achieve the same worthy end. That there are multiple-option cases where the relevant reasons are imperfect reasons is, I take it, uncontroversial. But if the existence of multiple-option cases is as uncontroversial as I claim, we might wonder why philosophers such as Raz (1999) and Gert (2003) have held that we cannot account for the basic belief simply by appealing to them. The answer is that they each believe that multiple-option cases are insufficiently numerous to account for the basic belief.

To demonstrate this, Raz cites a case in which a woman named Mary has the opportunity to see a powerful performance of a good play at her local theater tonight but decides to stay home instead (1999, p. 99). Raz's view is that both alternatives—going to the theater and staying home—are rationally permissible. And although he admits that many reasons are imperfect reasons (or what he calls reasons that “are not time-specific”), he denies that such reasons allow us to account for the basic belief. He says:

In many cases the reasons for doing one thing or another are not time-specific: the same reasons and the same opportunity to conform to them will apply on a number, sometimes an indefinite number, of occasions. This does not, however, explain the basic belief. Quite apart from the fact that delay is not costless, the basic belief applies to time-specific reasons as well. Mary, in our example, does not have to go to the play even on the last evening of its run. She may still just not feel like it and do something else instead. (1999, p. 100)

Likewise, Gert believes that the set of cases in which there are rational options is significantly broader than the set of multiple-option cases. Gert believes, for instance, that it is rationally permissible both to sacrifice \$200 to prevent forty children from suffering from serious malnutrition for forty days and to refuse to do so, choosing instead to spend that money on oneself (2003, pp. 8-9). As he sees it, this is a case in which there is a rational option, but he denies that it is a multiple-option case. Gert says:

Multiple-option cases depend crucially on the fact that the class of rationally justified options in any given case can plausibly be seen as ways of doing the *same thing*. In the examples that were used to motivate the justifying/requiring distinction, on the other hand, the pairs of justified options were always of the following form.

- (a) Make a sacrifice for an altruistic reason.
- (b) Do not make the sacrifice.

Such pairs cannot plausibly be construed as alternate ways of pursuing the same end or acting on the same reasons. (2003, p. 13)<sup>16</sup>

This, I will argue, is mistaken. I will show that options *a* and *b* can be construed as two alternative ways of pursuing the same end and of acting on the same reason.

As Gert admits, it is crucial to get the level of description right when assessing the rational status of an alternative, for whether it's true that I am doing what I am rationally required to be doing depends on how we describe what it is that I am doing (2003, pp. 13-14). To illustrate, consider again the case where I must get to the airport, and let us assume that I happen to be taking a taxi to get there. If we ask whether I'm rationally required to be doing *what I am doing*, the answer will depend on the level of description. If, on the one hand, we describe *what I am doing* as taking a taxi to the airport, the answer will be "No," because taking the bus is an equally attractive means of getting to the airport and I am rationally permitted to do that instead. If, on the other hand, we describe *what I am doing*, more generally, as going to the airport, the answer will be "Yes," because I am rationally required to go to the airport (or so we're supposing). So, at one level of description, I am rationally required to be doing *what I am doing*, and, at another, I'm not.

The same can be said of Gert's example (and, as we'll later see, of Raz's example as well). At a more general level of description, there are two courses of action that I might take over time, one that includes my acting self-interestedly now and altruistically later and another that includes my acting altruistically now and self-interestedly later, and it is plausible to construe these two courses of action as two ways of trying to achieve the same thing: a reasonable balance between altruistic acts and self-interested acts over time. If we think that the order in which I perform such acts is unimportant and that what's important is only that I end up striking a reasonable balance between the two types of acts, then it doesn't matter whether I act self-interestedly now and altruistically later or vice versa so long as, in the end, I perform a sufficient number of each over time.

In Raz's case, we can say something similar. If we acknowledge that Mary conceives of her agency as being extended over time, then we should think that, at the relevant level of description, the pertinent choice is not between seeing this play on its last night and staying home, but between two courses of action, containing both a certain amount of relaxation and a

<sup>16</sup> The letters *a* and *b* are not italicized in the original.

certain amount of cultural enrichment. So even though this is a good play and it's Mary's last opportunity to see it, this will not be Mary's last opportunity to do something as entertaining and culturally enriching as seeing this play. And surely she isn't rationally required to take advantage of every such opportunity. On the broader view, then, what each of us, including Mary, should be trying to do is to strike a reasonable balance between various worthy pursuits, such as relaxation and cultural enrichment, over time. It's important to note, then, that if Mary stays home tonight and relaxes, there will be other opportunities for her to similarly enrich herself, and that, if she goes to the play, there will be other opportunities for her to get the relaxation that she periodically needs. What's important, then, is not that she takes advantage of any specific opportunity for relaxation or cultural enrichment, but that she takes advantage of both sorts of opportunities sufficiently often over time.

To illustrate, suppose that Mary decides to stay home tonight while intending to take advantage of some future opportunity to do something culturally enriching. If we then ask whether she is rationally required to be doing *what she is doing*, the answer will depend, as it did above, on the level of description. If, on the one hand, we describe *what she is doing* as staying home tonight while intending to do something culturally enriching on some future night, the answer will be "No," because she is permitted instead to go to the theater tonight while intending to relax at home on some future night, which is an equally attractive means of striking a reasonable balance between relaxation and cultural enrichment over time. If, on the other hand, we describe *what she is doing*, more generally, as implementing a plan to strike some reasonable balance between relaxation and cultural enrichment over time, then the answer will be "Yes," because she is rationally required to strike such a balance over time. So, at one level of description, she is rationally required to be doing *what she is doing*, and, at another, she is not. So it seems that if we view Mary's options from a suitably general level of description, we find that this is a multiple-option case after all.

So while admitting that there are multiple-option cases, Raz and Gert claim that such cases are insufficiently numerous to account for the basic belief, and so they resort to more controversial claims in trying to account for the basic belief: in Raz's case, to the claim that there are widespread incommensurabilities among reasons, and, in Gert's case, to the claim that reasons have two separable dimensions of strength (i.e., requiring strength and justifying strength). I, however, will argue that multiple-option cases are sufficiently numerous to account for the basic belief and, thus, that we

needn't appeal to such controversial claims in accounting for the basic belief. My argument will be that Raz and Gert fail to appreciate the fact that most typical choice situations are multiple-option cases, because they fail to view our choices from the appropriate level of description: that of choosing between various courses of action over time. To be clear, I am not going to argue that the contentious claims that Raz and Gert appeal to in accounting for the basic belief are false or even that they are not needed to account for something else. I will argue only that they are not needed to account for the basic belief and, thus, that we can account for the basic belief with only the theoretical apparatus that all sides agree on: imperfect reasons and multiple-option cases. Also, I should warn the reader that I will not be defending the basic belief. As Raz pointed out, our task is to explain it, and that's what I aim to do and with as few assumptions as possible.

In the next section, I present my own theory of objective rationality—viz. securitism. I then show, in the subsequent section, how this theory enables us to account for the basic belief. In the last section, I explore securitism's suppositions and implications, defending it against various potential objections.

### §6.3 Securitism

§§ 6.3.1 *Objective rationality*: Before I state my theory, let me be clear on what it is a theory about. It is a theory about what it is *objectively* rational to do. Objective rationality is about what agents ought to do in the most fundamental and unqualified sense. So it's not about what they morally ought to do, nor is it about what they prudentially ought to do. Instead, it is about what they objectively ought to do, all things considered—see §1.4. Thus, to say that an act is objectively *irrational* is to say that the “action absolutely should not be performed” (GERT 2004, p. 137). Some suggest that we should understand objective rationality in terms of advisability: objectively irrational actions are those that no fully-informed and well-meaning adviser would recommend performing on the assumption that the agent would take the advice.<sup>17</sup> More precisely, though, we should

<sup>17</sup> See GERT 2004 (p. 140) for more on this idea. I add the qualification “on the assumption that the agent will take the advice” to eliminate the worry that in some cases the agent might be predisposed to rebel against such advice. I should also add that the advice needs to be complete. That is, it won't do to advise an agent merely to do *x* if doing *x* without simultaneously intending or doing *y* would be disastrous. What's more, we must realize that an agent's reasons for doing *x* could potentially diverge from an advisor's reasons for advising her to do *x*, and the objective rationality of an agent's doing *x* is a function of her

understand the notion of objective rationality in terms of reasons: an act is objectively irrational if and only if there are decisive reasons for the agent to refrain from performing it—that is, if and only if the agent objectively ought to refrain from performing it (see §1.4).<sup>18</sup> The objective rational status of an act is, then, purely a function of the reasons there are for and against performing it, irrespective of whether or not the agent is aware of them.

The subjective rational status of an act, by contrast, depends not on what reasons there are, but on what reasons the agent takes there to be or, alternatively, on the practical mental functioning of the agent. On the former view, an act is subjectively irrational if and only if the agent has beliefs whose truth would give her decisive reasons to refrain from performing it. And when the agent has inconsistent beliefs, the act will be subjectively rational relative to some beliefs, but subjectively irrational relative to others (PARFIT 2008). On the latter view, an act is subjectively irrational if and only if it indicates some failure in the practical mental functioning of the agent (GERT 2004, p. 160). I won't take a stand on which, if either, account of subjective rationality is correct. For my purposes, it is important only that we have a clear understanding of objective rationality, of which I now offer a theory.

Let me warn the reader, though, that the theory that I'm about to offer is unusual in a number of ways. So there will be a fair bit of explanation involved, and it won't be until the next section that I will explore the theory's implications with regard to the basic belief.

§§ 6.3.2 *Doing the best we can*: According to many normative theories, we are to assess the normative status of an action by comparing it to its alternatives. If we accept such a theory of objective rationality, we would assess the objective rational status of an action by weighing how much reason there is to perform it as opposed to its alternatives. More specifically, we might hold that an act is rationally permissible if and only if there is no alternative that there is even more reason to perform. Such a

reasons for doing  $x$ , not the advisor's reasons for advising her to do  $x$ . At best, then, the advice model is just an often useful (but also sometimes misleading) heuristic for getting a grip on our intuitions about what is objectively rational. It must, however, be used cautiously. For more on this issue, see §1.4.

<sup>18</sup> Another helpful way of understanding this notion, which was suggested to me by Peter de Marneffe, is as follows: an act is objectively irrational if and only if the set of true propositions provide decisive reasons for the agent to refrain from performing it. I'm assuming, here, that a reason for  $S$  to  $\varphi$  is just some fact that counts in favor of  $S$ 's  $\varphi$ -ing. Thus, there can be reasons for  $S$  to  $\varphi$  of which  $S$  is completely unaware, for  $S$  can be completely unaware that the relevant reason-providing facts obtain.



theory is, I think, quite promising. In any case, it's the sort of theory that I'll explore here.

Once we accept such a theory, the first question to ask is: Which are the relevant alternatives, such that one ought to perform the best alternative—i.e., the one that there is most reason to perform? One common answer is that the relevant alternatives are those that the agent *can* perform. Thus, a number of philosophers have claimed that we ought to do the best we can.<sup>19</sup> Let's call the alternatives that an agent, S, can perform 'personally possible' (ZIMMERMAN 1996, p. 46). The idea that S ought to do the best that's personally possible for her is plausible only if 'ought' implies 'personally possible' (i.e., 'can') and nothing more restrictive than that. For suppose that 'ought' implies not only 'personally possible', but also 'F' ('F' standing for some predicate), where the set of acts that are F is a proper subset of those that are personally possible. If such is the case, then there would be some acts that are personally possible for S but which S cannot be obligated to perform: those that are personally possible for S but not F. Thus, if there is such an F, we must reject the idea that agents ought to do the best that they can. I think that there is such an F and that 'F' stands for 'securable'. Let me explain.

To illustrate the difference between what's securable by an agent and what's personally possible for that agent, consider the following. Suppose that I'm on a low fat, low carbohydrate diet. It's now 3 p.m., and I'm thinking that I should grill up some low calorie vegetables and a lean boneless, skinless chicken breast for tonight's dinner. If this is going to be what I'm having for dinner, I'll need to get into my car, drive to the grocery store, buy the necessary ingredients, drive back home, marinate the chicken, cut up the vegetables, fire up the grill, wait for it to heat up, grill the chicken and vegetables, serve them, and eat them. But suppose that, as a matter of fact, no matter what, at 2 p.m., I plan or intend on doing later, I will not end up eating grilled chicken and vegetables for tonight's dinner. For, as a matter of fact, I'm going to end up eating pizza instead.<sup>20</sup> What's going to happen is this: while in the midst of preparing the chicken and vegetables, I'm going to get very hungry. And, in this state, I'm going to be

<sup>19</sup> See, for instance, FELDMAN 1986 and ZIMMERMAN 1996.

<sup>20</sup> Here and elsewhere, I'll be assuming that, for each set of actions that S might perform, there is some determinate fact as to what the world would be like were S to perform that set of actions. And I'm also assuming that, for each set of intentions that S might form, there is some determinate fact as to what the world would be like were S to form those intentions. This assumption is sometimes called *counterfactual determinism*—see, e.g., BYKVIST 2003, p. 30.

overcome by the temptation to go off my diet and will at 5 p.m. throw the chicken and vegetables back in the fridge, turn off the grill, and order pizza. It's not that I couldn't grill and eat the chicken and vegetables. Indeed, I could in that, were I to intend at 5 p.m. to continue with my plan to grill and eat chicken and vegetables, that's exactly what I would do. It's just that, as a matter of fact, I will at 5 p.m. abandon my plan and order pizza instead.

In this case, my eating chicken and vegetables at 6 p.m. is, as of 2 p.m., personally possible for me, but it is *not* something that is, as of 2 p.m., securable by me. It's personally possible for me, for there is a series of steps that I could take that culminate in my eating grilled chicken and vegetables at 6 p.m. tonight, and the following is true of each one of these steps: having taken the previous step or steps, I could take the next step in that, were I to intend to take the next step, I would succeed in doing so.

Nevertheless, my eating chicken and vegetables at 6 p.m. is not, as of 2 p.m., securable by me, for, as we're supposing, no matter what I intend to do now at 2 p.m., I will not end up eating grilled chicken and vegetables at 6 p.m. Even if I were, at 2 p.m., to form the most resolute intention to stick with my plan, it's a fact about my psychology that I would abandon it in the face of the ensuing temptation to eat pizza. Thus, there's absolutely nothing that I can intend to do now that will result in my eating chicken and vegetables at 6 p.m.—or so we're assuming. Of course, there are things that I could intend to do now that would prevent me from eating pizza at 6 p.m.. I could, for instance, intend now to commit some crime in front of a police officer so that I'll end up in jail by 6 p.m. Nevertheless, there is nothing that I could intend to do now at 2 p.m. that would ensure that I end up eating grilled chicken and vegetables at 6 p.m. Thus, my sitting in jail at 6 p.m. is, as of 2 p.m., securable by me, but my eating chicken and vegetables at 6 p.m. is not.

With these examples in mind, I can now define the two notions more precisely. Let ' $\alpha_i$ ' and ' $\alpha_j$ ' be variables that range over sets of actions. The acts in these sets may be simple or compound, synchronous or asynchronous, consecutive or inconsecutive. The definitions of 'securable' and 'personally possible' are, respectively, as follows:

- 6.1 A set of actions,  $\alpha_j$ , beginning at  $t_j$  is, as of  $t_i$  ( $t_i \leq t_j$ ), *securable* by S if and only if there is a set of actions,  $\alpha_i$  ( $\alpha_i \subseteq \alpha_j$ ), such that both of the following hold: (1) S would perform  $\alpha_j$  if S were, at  $t_i$ , to

intend to perform  $\alpha_i$ , and (2) S is not, at  $t_i$ , in a state that's incompatible with S's intending to perform  $\alpha_i$ .<sup>21</sup>

- 6.2** A set of actions,  $\alpha_j$ , consisting in a number of steps extending from  $t_j - t_n$  is, as of  $t_i$ , *personally possible* for S if and only if S's performing the first step at  $t_j$  is, as of  $t_i$ , securable by S, and if S were to perform the first step at  $t_j$ , S's performing the second step at  $t_k$  would be, as of  $t_j$ , securable by S, and if S were to perform the first two steps at  $t_j$  and  $t_k$ , respectively, S's performing the third step at  $t_l$  would be, as of  $t_k$ , securable by S, and so forth until all the steps have been completed ( $t_i < t_j < t_k < t_l \dots < t_n$ ).<sup>22</sup>

Both definitions need clarifying. I'll start with 6.1. First, note that I'm assuming that 'intentions' should be construed broadly so as to include plans, policies, resolutions, and the like. We should construe intentions in this way, for we are planning agents. We conceive of our agency as being extended over time, and because of this, we often plan, undertake, and then, finally, complete temporally extended activities and projects (BRATMAN 2007, p. 21). We do not simply act from moment to moment. Instead, we form future-directed intentions and adopt both plans and policies, which help us to organize and coordinate our activities over time (BRATMAN 2007, p. 26).

As Michael Bratman explains, "[p]lans typically concern relatively specific courses of action extended over time" (2007, p. 27). For instance, one might plan to spend the morning grading exams and the afternoon playing golf. Policies, by contrast, involve a general commitment "to a certain kind of action on certain kinds of potentially recurrent occasions—for example, to buckling up one's seat belt when one drives, or to having at most one beer at dinner" (BRATMAN 2007, p. 27). Besides plans and policies, there are also resolutions. A resolution is an intention that one has resolved not to reconsider even in the face of anticipated temptation to do

<sup>21</sup> Perhaps, I should replace condition 1 with something stronger: with, say, (1') S would perform  $\alpha_j$  if and only if S were, at  $t_i$ , to intend to perform  $\alpha_i$ . I'm uncertain as to which of condition 1 and condition 1' to accept. They do, however, have different implications in the following sort of case. Suppose that I will perform  $\alpha_j$  no matter what, if anything, I intend at  $t_i$  to do. On condition 1, but not on condition 1',  $\alpha_j$  would count as being, as of  $t_i$ , securable by S. I'm inclined to go with condition 1, but if the reader prefers something stronger along the lines of condition 1', I'm not necessarily opposed to substituting condition 1' for condition 1.

<sup>22</sup> This is inspired by similar definitions by both Goldman (1978, p. 193) and Zimmerman (1996, p. 46)

so (HOLTON 2009, pp. 9-12).<sup>23</sup> Thus, someone might resolve to exercise or to refrain from eating sweets.

Second, to say that S performs a set of actions,  $\alpha_j$ , is just to say that S performs each act in that set. Suppose, for instance, that I plan to spend this morning working on this chapter and tomorrow morning grading exams. I perform this set of actions if and only if I spend this morning working on this chapter and tomorrow morning grading exams.

Third, condition 2 is needed given the following sort of case. Suppose that S is at  $t_i$  in a coma. Presumably, given S's unconscious state at  $t_i$ , no set of actions is, as of  $t_i$ , securable by S. Yet the following counterfactual may be true: S would perform  $\alpha_j$  if S were, at  $t_i$ , to intend to perform  $\alpha_i$ . That is, in the closest possible world in which S intends at  $t_i$  to perform  $\alpha_i$  (a world in which S is conscious at  $t_i$ ), S performs  $\alpha_j$ . Thus, we must add: (2) S is not, at  $t_i$ , in a state that's incompatible with S's intending to perform  $\alpha_i$ . Adding this condition allows us to account for the fact that a set of actions won't be, as of  $t_i$ , securable by S unless S is, at  $t_i$ , in a state that's compatible with S's having or forming the relevant intentions. And, clearly, being in a coma is not a state that's compatible with having and forming the relevant intentions.

Fourth, it's important to note that the relevant intention here is not the intention to perform  $\alpha_j$ , but the intention to perform  $\alpha_i$ . Thus, 6.1 allows that a set of actions,  $\alpha_j$ , can be, as of  $t_i$ , securable by S even though S is, at  $t_i$ , in a state that's incompatible with S's intending to perform  $\alpha_j$ . To illustrate, consider Sam, a twelve-year-old boy who, on his twelfth birthday (more specifically, at  $t_i$ ), decides that he wants to earn a PhD in mathematics. And let's assume that  $\alpha_j$  is the set of all the acts that Sam would need to perform in order to achieve this goal. Now, at  $t_i$ , Sam certainly doesn't know what all he needs to do to earn a PhD in mathematics. For instance, he doesn't know that he is going to need to prove Gödel's incompleteness theorems on his comprehensive exams. Indeed, as of  $t_i$ , he hasn't even heard of Gödel's incompleteness theorems. Given this, it is, as of  $t_i$ , psychologically impossible for Sam to form the intention to perform  $\alpha_j$ . For one,  $\alpha_j$  includes acts that Sam can't at  $t_i$  intend to perform, simply because Sam doesn't at  $t_i$  have the requisite concepts. Sam can't, for instance, intend to prove Gödel's incompleteness theorems when he doesn't even have the concept of a theorem or the concept of a formal system's being incomplete. For another,

<sup>23</sup> On one plausible view, a resolution involves both a first-order intention to perform a certain action and a second-order intention not to let that first-order intention be deflected by anticipated contrary inclinations—see HOLTON 2009, pp. 11-12.

$\alpha_j$  is simply too large and too complex for Sam to hold consciously before his mind in the way that is necessary for intending to perform  $\alpha_j$ . Thus, Sam is, at  $t_i$ , in a state that's incompatible with his intending to perform  $\alpha_j$ .

Nevertheless, 6.1 allows that  $\alpha_j$  can be, as of  $t_i$ , securable by Sam. It will be securable so long as there is some other set of actions,  $\alpha_i$  ( $\alpha_i \subseteq \alpha_j$ ), such that both of the following hold: (1) S would perform  $\alpha_j$  if S were, at  $t_i$ , to intend to perform  $\alpha_i$ , and (2) S is not, at  $t_i$ , in a state that's incompatible with S's intending to perform  $\alpha_i$ . Suppose, for instance, that the intention to perform  $\alpha_i$  simply consists in the intention to earn a PhD in mathematics and that Sam is in a state that's compatible with his forming this intention. And let's assume that if Sam were, at  $t_i$ , to intend to earn a PhD in mathematics, he would perform the very large and complex set of actions that is required to do so, including, let's assume, proving Gödel's incompleteness theorems. In that case, 6.1 implies that Sam's performing  $\alpha_j$  is, as of  $t_i$ , securable by him even though he is at  $t_i$  in a state (that of having both limited mental abilities and limited conceptual apparatus) that's incompatible with his intending to perform  $\alpha_j$ .

With regard to 6.2, I should clarify what I mean by 'the first step', 'the second step', 'the third step', etc. The first step in a set of actions,  $\alpha_j$ , consisting in a number of steps extending from  $t_j - t_n$ , is the act (possibly a synchronous compound act) that would be performed first at  $t_j$ . The second step would be the act that comes next, temporally speaking, in that set, and so on and so forth for each of the steps. So, for instance, if  $\alpha_j$  consists in accelerating while changing lanes at  $t_1$ , exiting the freeway at  $t_4$ , and pulling into a gas station at  $t_6$ , then step one consists in accelerating while changing lanes at  $t_1$ , step two consists in exiting the freeway at  $t_4$ , and step three consists in pulling into a gas station at  $t_6$ .

The main difference between those sets of actions that are securable and those that are merely personally possible is that the latter class is often considerably larger than the former. For although every set of actions that is, as of  $t_i$ , securable by S will also be, as of  $t_i$ , personally possible for S, the converse doesn't hold. Some sets of actions that are not, as of  $t_i$ , securable by S will be, as of  $t_i$ , personally possible for S.

§§ 6.3.3 'Obligatory' implies 'securable': As I said above, I think 'ought' implies not only 'can', but also 'securable'. More precisely, I think that S's being objectively obligated, as of  $t_i$ , to  $\varphi$  implies not only that S's  $\varphi$ -ing is, as of  $t_i$ , personally possible for S, but also that S's  $\varphi$ -ing is, as of  $t_i$ , securable by S. Why do I think this? There are two arguments that I find persuasive.

The first of these I owe, in broad strokes, to Richard Chappell (2009). It begins with the plausible assumption that an agent can't rationally intend at  $t_i$  to  $\varphi$  if she knows that she won't  $\varphi$  even if she intends at  $t_i$  to  $\varphi$ .<sup>24</sup> Now, any act that is not, as of  $t_i$ , securable by S is one that S won't perform even if she intends at  $t_i$  to perform it. This just follows from the definition of 'securable'. So, given our initial assumption and the definition of 'securable', it follows that any agent who is, at  $t_i$ , both perfectly rational and fully informed about her present alternatives will be unable to intend at  $t_i$  to perform any act that is not, as of  $t_i$ , securable by her. For being fully informed about her present alternatives she would know that such an intention would be ineffective, and, knowing this and being perfectly rational, she would be incapable of forming that intention at  $t_i$ .<sup>25</sup> So unless we want to hold that an agent can, as of  $t_i$ , be objectively rationally required to perform an act that she can't rationally intend at  $t_i$  to perform when fully informed about her present alternatives, we must deny that an agent can, as of  $t_i$ , be objectively rationally required to perform an act that is not, as of  $t_i$ , securable by her. And, of course, we must deny that an agent can, as of  $t_i$ , be objectively rationally required to perform an act that she can't rationally intend at  $t_i$  to perform when fully informed about her present alternatives, for an act that is, as of  $t_i$ , *objectively* rationally required of her is just one that she would intend at  $t_i$  to perform if she were, at  $t_i$ , both perfectly rational and fully informed about her present alternatives (see §1.4). Therefore, we must deny that an agent can, as of  $t_i$ , be objectively rationally required to perform an act that is not, as of  $t_i$ , securable by her.

To illustrate, take the case from above, except assume that I *know* that, no matter what I intend to do now at 2 p.m., I will not end up eating grilled chicken and vegetables at 6 p.m. Knowing this, I don't see how I could rationally intend at 2 p.m. to eat chicken and vegetables at 6 p.m. And,

<sup>24</sup> Many philosophers have endorsed either this thesis or some stronger thesis that entails this one. J. H. Sobel claims that "a person cannot intend an action she is sure that she will not do" (1994, p. 239). Wallace claims that "the intention to do  $x$  requires at least the belief that it is *possible* that one do  $x$ " (2001, p. 20). Bratman claims that "there will normally be irrationality in intending to  $A$  and believing that one will not  $A$ " (1987, p. 38). Harman (1976, p. 432) and Velleman (1989) have each claimed that intending to  $\varphi$  entails the belief that one will  $\varphi$ .

<sup>25</sup> I mean for her present alternatives to include not only various alternative actions, but also various alternative intentions. Note, though, that to say that S is, at  $t_i$ , both perfectly rational and fully informed about her present alternatives does not imply that S has been perfectly rational in the past or will be perfectly rational in the future. Indeed, part of what is entailed in S's being, at  $t_i$ , fully informed about her present alternatives is knowing what she will do at future times given her predictable lack of perfect rationality at those future times.

given this, I don't see how we can rightfully claim that I am, as of 2 p.m., objectively rationally required to eat grilled chicken and vegetables at 6 p.m. How can I be objectively rationally required to do what I can't even rationally intend to do when I am fully informed about my present alternatives? It's true that my eating grilled chicken and vegetables at 6 p.m. is, as of 2 p.m., personally possible for me, but it is not, as of 2 p.m., securable by me. There is nothing that I can do or intend to do at 2 p.m. that will make it the case that I eat grilled chicken and vegetables at 6 p.m. Indeed, the situation is worse: no matter what I do or intend to do at 2 p.m., I will not eat grilled chicken and vegetables at 6 p.m. But, then, how can I be, as of 2 p.m., under an objective obligation to eat grilled chicken and vegetables at 6 p.m. when, knowing this, rationality prohibits me from intending at 2 p.m. to eat grilled chicken and vegetables at 6 p.m.?

The second argument for the claim that *S*'s being objectively obligated, as of  $t_i$ , to  $\varphi$  implies that *S*'s  $\varphi$ -ing is, as of  $t_i$ , securable by *S* comes from Holly S. Goldman (now Holly M. Smith)—see GOLDMAN 1978, pp. 194-195. She argues that the range of alternatives that *S* can be obligated, as of  $t_i$ , to perform can be no wider than those which *S* can, as of  $t_i$ , secure. Otherwise, we would have to allow that *S* can, as of  $t_i$ , be obligated to  $\varphi$  at  $t_j$  even though there is no way, as of  $t_i$ , for *S* to ensure that this obligation is fulfilled. And what would be the point in claiming that *S* is, as of  $t_i$ , obligated to  $\varphi$  at  $t_j$  if she can't, at  $t_i$ , either effectively intend to  $\varphi$  at  $t_j$  or effectively intend to do anything else that would result in her  $\varphi$ -ing at  $t_j$ ? It seems that such a prescription would be utterly pointless, for not even an agent who was, at  $t_i$ , both perfectly rational and fully informed about her present alternatives could make any practical use of it.

If no present intention of *S*'s could lead *S* to perform the required action, then no present knowledge on *S*'s part of that putative requirement could lead to its being fulfilled (GOLDMAN 1978, p. 195). And if no present knowledge of the requirement could lead even the most conscientious agent to fulfill it, then there is absolutely no point in ascribing that obligation. And if there's no point in ascribing it, then we should deny that it exists. For an objective obligation, as of  $t_i$ , to  $\varphi$  must be something that an agent who was, at  $t_i$ , both perfectly rational and fully informed about her present alternatives could use for practical guidance. That is, there has to be something that such an agent could, at  $t_i$ , do or intend to do that would ensure that the putative obligation would be fulfilled.

To illustrate, consider again the case from above, and again assume that I know that no matter what I intend to do now at 2 p.m., I will not end up eating grilled chicken and vegetables at 6 p.m. Let's further assume that

I am, at 2 p.m., perfectly rational. Now suppose that you tell me that I am, as of 2 p.m., objectively rationally required to eat grilled chicken and vegetables at 6 p.m. What am I supposed to do with this “information”? Given that my eating grilled chicken and vegetables at 6 p.m. is not, as of 2 p.m., securable by me, there’s nothing that I can do with it, for there is no present action or intention that would lead me to fulfill this putative obligation. So it seems utterly pointless. And, again, it seems that we should deny that there are any such utterly pointless obligations. “Obligations” that would be of no practical use even to an agent who was, at the time, both perfectly rational and fully informed about her present alternatives are not genuine obligations.

§§ 6.3.4 *Securitism*: Given the above arguments, we should hold that  $S$ ’s being objectively obligated, as of  $t_i$ , to  $\varphi$  implies that  $S$ ’s  $\varphi$ -ing is, as of  $t_i$ , securable by  $S$ . And, thus, we need a theory of objective rationality that accounts for this fact. I hope to offer such a theory, which I call *securitism*. But before I can state the theory, I need to add a couple more definitions to those given above:

- 6.3 One set of actions,  $\alpha_j$ , *includes* another,  $\alpha_i$ , if and only if every element in  $\alpha_i$  is also an element in  $\alpha_j$ —that is, if and only if  $\alpha_i \subseteq \alpha_j$ .
- 6.4 A set of actions,  $\alpha_j$ , that is, as of  $t_i$ , securable by  $S$  is a *maximal set* if and only if there is no set of actions that is, as of  $t_i$ , securable by  $S$  that includes  $\alpha_j$  as a proper subset.<sup>26</sup>

I can now state securitism as follows:

- SEC** It is, as of  $t_i$ , objectively rationally permissible for  $S$  to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively rationally permissible maximal sets of actions that are, as of  $t_i$ , securable by  $S$  includes  $S$ ’s performing  $\alpha_j$ .<sup>27</sup>

<sup>26</sup> One set of actions,  $\alpha_j$ , is a proper subset of another,  $\alpha_i$ , if and only if every element in  $\alpha_j$  is also an element in  $\alpha_i$  but not every element in  $\alpha_i$  is an element in  $\alpha_j$ —that is, if and only if  $\alpha_j \subset \alpha_i$ .

<sup>27</sup> As mentioned in note 20, I’m assuming counterfactual determinism. If this assumption is false, we’ll need to correlate each set of intentions with a probability distribution over the sets of possible worlds that might be actualized if  $S$  were to form those intentions, and we’ll need to replace securitism with something along the following lines: It is objectively



This theory has two rather unusual features. First, it tells us to apply normative principles directly only to maximal sets of actions. We, then, assess non-maximal sets of actions indirectly in terms of whether they are included in some permissible maximal set of actions. Second, securitism relativizes permissions—and, consequently, obligations—to times.<sup>28</sup> I'll explain each of these features below. But, before I do, it's important to note that securitism is not itself a substantive view, for it doesn't tell us how to assess maximal sets of actions (hereafter: 'MSAs'). Securitism is, for instance, neutral with regard to whether the correct normative principles are teleological or non-teleological. Of course, given my arguments in Chapter 3, I would favor a teleological version of securitism, where it is held that a given maximal set of actions,  $MSA_i$ , is, as of  $t_i$ , objectively rationally permissible if and only if, and because, its outcome is not, on  $S$ 's evaluative ranking, outranked by that of any other MSA that is, as of  $t_i$ , securable by  $S$ .<sup>29</sup> But, for the purposes of this chapter, I will pretend to be neutral as to how MSAs are to be assessed.

Focusing, then, on securitism in general and not on any particular substantive version of securitism, I'll still need to explain its two unusual features, starting with the first one, which is that it holds that normative

rationally permissible for  $S$  to intend at  $t_i$  to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, there is some set of intentions that  $S$  can form at  $t_i$  that is optimal in terms of expected deontic value and that includes  $S$ 's intending at  $t_i$  to perform  $\alpha_j$ . The deontic value of  $S$ 's intending to  $\varphi$  is a measure of how much reason there is for  $S$  to perform  $\varphi$ . The expected deontic value of  $S$ 's intending to  $\varphi$  is calculated by multiplying the deontic value of  $S$ 's intending to  $\varphi$  by the probability that  $S$  would  $\varphi$  were  $S$  to intend to  $\varphi$ . Of course, in order for this to serve as a suitable replacement for securitism, we would need to supplement it with an account of what it is objectively rationally permissible *to do* given this account of what it is objectively rationally permissible *to intend to do*. Unfortunately, I have no such account to offer at the moment. (I thank Andy Specht for helpful discussions of these issues.)

Securitism is in many ways inspired by Goldman's Principle 4—see 1978, p. 202.

<sup>28</sup> An act is obligatory if and only if it is the only permissible alternative.

<sup>29</sup> More specifically, I endorse *teleological maximizing securitism* (or telemax securitism): (a) It is, as of  $t_i$ , objectively rationally permissible for  $S$  to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively rationally permissible maximal sets of actions that are, as of  $t_i$ , securable by  $S$  includes  $S$ 's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively rationally permissible if and only if, and because, its outcome is not, on  $S$ 's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by  $S$ . (c)  $MSA_i$ 's outcome is not, on  $S$ 's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by  $S$  if and only if, and because, there is no alternative MSA that is, as to  $t_i$ , securable by  $S$  whose outcome  $S$  has more reason, all things considered, to want to obtain.

principles are to be applied directly only to MSAs. To understand why securitism has this unusual feature, we must first understand why it would be inadequate for a normative theory to assess only *minimal acts*. Minimal acts are the most basic acts. They are the indivisible atoms of action. Most actions are not minimal acts, as most acts are composed of a number of other more basic acts. For instance, the act of firing a semi-automatic handgun is a compound act that is comprised of a sequence of more basic actions: inserting the magazine, releasing the slide, taking off the safety, taking aim, and squeezing the trigger. Minimal actions, by contrast, do not consist of other more basic acts. As Jordan Howard Sobel explains,

*A minimal action is one which, once begun, cannot be stopped by its agent short of completion: once initiated a minimal action is to this extent out of an agent's control. All instantaneous actions are necessarily minimal: for example, ...placing a bet. And some non-instantaneous actions are contingently minimal: for example, ...beheading by guillotine. (1976, p. 198)*

It's fairly obvious that a normative theory must do more than just assess minimal acts. Most of our practical deliberations concern non-minimal acts. For instance, an agent might deliberate about whether to fire a semi-automatic handgun. Or she might deliberate about whether to accelerate while turning right. Or she might deliberate about whether to grade this morning and mow the lawn tomorrow morning or vice versa. And most of our intentions are intentions to perform non-minimal acts. For instance, I intend to turn on the alarm and then lock the door when I leave the house this morning. And many of our intentions take the form of plans that involve performing sets of actions. For instance, this Thursday, I plan on dropping off my daughter at school in the morning, returning home to pack, driving down to Tucson in the early afternoon, picking up my friend at the airport there, and then afterwards heading for the Westward Look Resort, which is where we'll both be attending a workshop. Since our deliberations and intentions often concern such sets of actions, a normative theory must be able to assess them and not just minimal acts.

There are three possible approaches to using normative principles to assess both minimal actions as well as sets of actions. These are:

*The Bottom-Up Approach:* Use normative principles to directly assess the permissibility of only minimal acts and then agglomerate to determine the permissibility of sets of action.

*The Across-the-Board Approach:* Use normative principles to directly assess the permissibility of minimal acts, maximal sets of actions, and everything in between.

*The Top-Down Approach:* Use normative principles to directly assess the permissibility of only maximal sets of actions and then distribute to determine the permissibility of non-maximal sets of actions, including those that contain only one minimal act.

Securitism employs the top-down approach. To defend securitism, then, I must explain why this approach is superior to the other two. I'll start by considering the bottom-up approach. The problem with this approach is that it must rely on the following implausible principle, which holds that *permissibility agglomerates over conjunction* (hence: PAC):

$$\text{PAC} \quad [P(S, t_i, x_1), P(S, t_i, x_2), \dots, \& P(S, t_i, x_n)] \rightarrow P(S, t_i, [x_1, x_2, \dots, \& x_n]).$$

Here, 'P(S,  $t_i$ ,  $\varphi$ )' stands for 'S is, as of  $t_i$ , permitted to  $\varphi$ '.

To see why PAC is implausible, we need to consider only one simple example, which I'll call *The Two Medicines*. Assume that Zeke is suffering from a potentially fatal disease. Fortunately, Zeke can cure himself either by taking medicine M1 at  $t_3$  (call this  $a_1$ ) or medicine M2 at  $t_5$  (call this  $a_2$ ). He must, however, be careful not to do both, for this would be a lethal combination. Here, then, is a case where, contrary to PAC, an agent is, as of  $t_1$ , permitted to perform  $a_1$  and is also, as of  $t_1$ , permitted to perform  $a_2$ , but is not, as of  $t_1$ , permitted to perform both  $a_1$  and  $a_2$ . Thus, we should reject PAC and the approach that presupposes it: viz., the bottom-up approach.

The across-the-board approach doesn't fair any better. The problem with this approach is that it commits us to denying the following plausible principle, which holds that *permissibility distributes over conjunction* (hence: PDC):

$$\text{PDC} \quad P(S, t_i, [x_1, x_2, \dots, \& x_n]) \rightarrow [P(S, t_i, x_1), P(S, t_i, x_2), \dots, \& P(S, t_i, x_n)].$$

To illustrate, consider the following egoistic normative principle:

$$\text{EGO} \quad \text{It is, as to } t_i, \text{ objectively rationally permissible for S to perform } \varphi \text{ if and only if, and because, S's performing } \varphi \text{ would}$$

produce more utility for S than any other personally possible alternative would.

If we adopt EGO and take the across-the-board approach, letting ' $\varphi$ ' range over minimal acts, maximal sets of actions, and everything in between, then we must reject PDC. To see why, consider the following case, which I base loosely on a case from JACKSON AND PARGETTER 1986 (p. 235):

*Professor Procrastinate:* Professor Procrastinate receives, at  $t_1$ , an invitation to review a book. He is the best person to do the review and has the time to do it. The best thing for him and for others would be if he accepts the invitation at  $t_2$  and then writes the review when the book arrives at  $t_4$ . The second best thing for him and for others would be if he declines the invitation at  $t_2$ . The worst thing for him and for others would be if he accepts the invitation at  $t_2$  and then never writes the review, and this is what would in fact happen were he to accept the invitation at  $t_2$ . Assume, then, that although Professor Procrastinate's accepting the invitation at  $t_2$  and then writing the review at  $t_4$  is, as of  $t_1$ , personally possible for him, it is not, as of  $t_1$ , securable by him. That is, there is, as of  $t_1$ , nothing that he can do or intend to do that would ensure that he accepts the invitation at  $t_2$  and then writes the review at  $t_4$ . No matter how resolute he is, at  $t_1$ , in his intention to write the review at  $t_4$ , he will end up forever procrastinating, thereby failing ever to write the review.

EGO implies that it is, as of  $t_1$ , permissible for Professor Procrastinate both to accept the invitation at  $t_2$  and to write the review at  $t_4$ , for performing this compound act would produce more utility for Professor Procrastinate than any other relevant alternative would. (Here, the relevant alternatives are: (1) both to accept the invitation at  $t_2$  and to refrain from writing the review at  $t_4$ ; (2) both to decline the invitation at  $t_2$  and to refrain from writing the review at  $t_4$ ; and (3) both to ignore the invitation at  $t_2$  and to refrain from writing the review at  $t_4$ —writing the review at  $t_4$  without accepting the invitation at  $t_2$  is not, I'm supposing, an option.) Yet, EGO also implies that it is, as of  $t_1$ , impermissible for Professor Procrastinate to accept the invitation at  $t_2$ , for doing so would produce less utility for Professor Procrastinate than the only relevant alternative (i.e., declining the invitation) would. These two implications are incompatible with PDC. Thus, adopting EGO while taking the across-the-board approach forces us to deny PDC. Of course, I've considered only one possible normative

principle, viz. EGO, but it should be fairly obvious how the result generalizes, for the only way that the top-down approach differs from the across-the-board approach is that the former, but not the latter, ensures that PDC will hold.

I find that denying PDC is high price to pay for adopting the across-the-board approach and so I think that we should reject this approach. But not everyone agrees with me. Some philosophers think that it's perfectly plausible to deny PDC. For instance, Jackson and Pargetter claim that there are obvious counterexamples to this sort of principle. The particular principle that they're concerned with is the one that says that 'ought' distributes over conjunction, but everything they say with respect to this principle applies to PDC, *mutatis mutandis*. They say,

Perhaps an overweight Smith ought to stop smoking and eat less, but it may not be true that he ought to stop smoking. For it may be that were he to stop smoking he would compensate by eating more. It is clear to many overweight smokers, and their doctors, that they ought to stop smoking and eat less, while it is far from clear to them that they ought to stop Smoking. (JACKSON AND PARGETTER 1986, p. 247)

Assuming that, typically, S ought to  $\varphi$  only if S ought also to intend to  $\varphi$  and that the case that we're dealing with here is a typical one, then Jackson and Pargetter's remarks imply both (i) that Smith ought, as of  $t_0$ , to intend both to stop smoking and to eat less and (ii) that Smith ought, as of  $t_0$ , to intend not to stop smoking.<sup>30</sup> But this can't be right, because there is a consistency requirement on intentions. "There is, in particular, a rational demand that one's intentions, taken together with one's beliefs, fit together into a consistent model of one's future" (BRATMAN 2009, p. 29). We can state this demand more precisely if we stipulate that when someone intends to  $\varphi$  the propositional content of her intention is the proposition that she  $\varphi$ s. The requirement, then, is that S ought rationally to be such that the set consisting of all the propositional contents of S's beliefs and all the propositional contents of S's intentions is logically consistent (ROSS 2009, p. 244). Now, clearly implications i and ii violate this requirement. An agent can't rationally intend not to stop smoking while also intending both to

<sup>30</sup> There may be cases where S ought to  $\varphi$  even though it is not the case that S ought to intend to  $\varphi$ . Perhaps, Sergio ought to act spontaneously, but he can't do so by intending to act spontaneously. Nevertheless, I don't think that it's plausible to suppose that smoking and eating less are like acting spontaneously in that they can only be done by not intending to do them. So we should think, for instance, that if Smith ought both to stop smoking and eat less, then he ought to intend both to stop smoking and to eat less.

stop smoking and eat less. It seems, then, that we should accept PDC, because if we were to deny PDC we would have to hold that agents are sometimes rationally permitted to have intentions with logically inconsistent propositional contents, which is irrational.

Jackson and Pargetter go wrong, I believe, in thinking that Smith ought, as of  $t_0$ , to stop smoking and eat less. If we assume, as they do, that were Smith to stop smoking he would compensate by eating more, then we should deny that Smith's stopping smoking and eating less is, as of  $t_0$ , securable by Smith. And since, as I argued above, 'ought' implies 'securable', we should further deny that their claim that Smith ought, as of  $t_0$ , to stop smoking and eat less. And if we do that, we find that this is no counterexample to PDC after all.

Now, the only way to ensure that PDC holds short of accepting the implausible bottom-up approach is to accept the top-down approach. This is the approach that securitism takes. In the case of Professor Procrastinate, securitism implies, given certain plausible assumptions about how to assess MSAs, that Professor Procrastinate should, as of  $t_1$ , decline the invitation. But it denies that Professor Procrastinate should, as of  $t_1$ , both accept the invitation and write the review. It denies this, because this is not, as of  $t_1$ , securable by Professor Procrastinate.

As mentioned above, securitism has a second unusual feature: it relativizes permissions—and, consequently, obligations—to times. This is necessary, because obligations can change over time. For instance, the following is possible: S was, as of  $t_1$ , obligated to perform  $x$  but is not, as of  $t_2$ , obligated to perform  $x$ . Obligations can change over time, because both what's personally possible and what's securable can change over time and 'obligatory' implies both 'personally possible' and 'securable'. To illustrate, consider the following:

Suppose that Smith ought to give her students a list of paper topics on Wednesday. If Smith's obligations do not change over time, then it is always the case that Smith ought to provide the list on Wednesday. Wednesday comes, however, and Smith arrives in class without the list. If 'ought' implies 'can', it then follows that Smith is able on Wednesday to give the class the list. Yet although this something that Smith was once able to do, she is no longer able to do it. (CURRAN 1995, p. 72)

So we either have to reject the plausible view that 'ought/obligation' implies 'can/personally possible' as well as 'securable' or accept that obligations and permissions need to be time-indexed. I prefer the latter.<sup>31</sup>

<sup>31</sup> See GOLDMAN 1976 for more on this issue.

Although securitism has some unusual features not found in other more traditional theories, I think that this is a merit of the view. Traditional theories often fail to account for the fact that our abilities and hence our obligations change over time. And they often fail to provide criteria for assessing sets of actions, despite the fact that we often deliberate about, and form intentions to perform, sets of actions. For these reasons, I think that securitism is a promising alternative to such traditional moral theories. It's also promising in that it can, as I'll show below, account for the basic belief.

#### §6.4 Securitism and the basic belief

Securitism allows us to account for the basic belief, for, on securitism, it turns out that most choice situations are multiple-option cases, cases where there is more than one way for the agent to ensure that she performs some rationally permissible MSA. Consider, for instance, that S will, on securitism, have the rational option to act either altruistically or self-interestedly at  $t_1$  provided that there is both a rationally permissible MSA that is, as of  $t_0$ , securable by S in which S acts altruistically at  $t_1$  and a rationally permissible MSA that is, as of  $t_0$ , securable by S in which S acts self-interestedly at  $t_1$ . Furthermore, S will, on securitism, have the rational option of performing any of numerous self-interested acts at  $t_1$  (e.g., watching TV, working on a book, going to the theater, etc.) provided that there is, in each case, a rationally permissible MSA that is, as of  $t_0$ , securable by S in which S performs that self-interested act at  $t_1$ . And the same goes for S's performing any of numerous altruistic acts at  $t_1$ .

But will these provisos hold? The answer, I believe, is 'yes'. We should think that, in most choice situations, there are rationally permissible MSAs in which S performs various different types of self-interested acts as well as rationally permissible MSAs in which S performs various different types of altruistic acts, for no matter what specific criteria we think should be used to assess the permissibility of MSAs, it's clear that in assessing MSAs it often doesn't matter whether the first act in that set is of one type or another so long as, in the end, the same balance among the various types is achieved.<sup>32</sup> Thus, the basic belief only seemed puzzling, because we were taking too narrow a view of our choices. If we consider my choice of what

<sup>32</sup> I admit that sometimes which is performed now will affect whether or not a reasonable balance will be achieved in the end, for sometimes it's true to say that, unless I do something now, it will never get done. In such instances, procrastination is irrational. I'll have more to say about this in §§ 6.5.4.

to do from a perspective that ignores the fact that my agency is extended over time, it does seem difficult to believe that I have just as much reason to watch TV as to volunteer for Oxfam or to work on this book. But when we take the broader view and compare MSAs that all involve the same (or roughly the same) proportion of relaxing acts, altruistic acts, and career-furthering acts, just in different temporal sequences, we see that they are all equally well supported by reason, for they each constitute an equally effectual means to my performing a rationally permissible MSA.<sup>33</sup> Thus, I have, on securitism, the rational option of performing any one of them now.

To take a more concrete example, suppose that I am, as of  $t_0$ , rationally permitted to perform either  $a_1$  or  $a_4$  at  $t_1$  but that I am, as of  $t_0$ , rationally required to perform one of these two at  $t_1$ . This means that if we were to look at all of the MSAs that are, as of  $t_0$ , securable by me, we would find (1) that there is, at least, one rationally permissible MSA in which I perform  $a_1$  at  $t_1$ , (2) that there is, at least, one rationally permissible MSA in which I perform  $a_4$  at  $t_1$ , and (3) that there is no rationally permissible MSA in which I perform neither  $a_1$  nor  $a_4$  at  $t_1$ . So my options are:

- (A) Undertake some rationally permissible MSA that includes my performing  $a_1$  at  $t_1$  and, thus, perform  $a_1$  at  $t_1$ .
- (B) Undertake some rationally permissible MSA that includes my performing  $a_4$  at  $t_1$  and, thus, perform  $a_4$  at  $t_1$ .
- (C) Undertake some rationally impermissible MSA that includes neither my performing  $a_1$  at  $t_1$  nor my performing  $a_4$  at  $t_1$  and, thus, perform neither  $a_1$  nor  $a_4$  at  $t_1$

This is clearly a multiple-option case, for there is more than one way for me to achieve the same end: that of undertaking some rationally permissible MSA. One way is by opting for A, and another is by opting for B. And the reason that I have to perform  $a_1$  (or  $a_4$ ) at  $t_1$  (viz., that doing so is a means to my performing a rationally permissible MSA) is an imperfect reason, for it is equally a reason to perform  $a_4$  (or  $a_1$ ) at  $t_1$ .

Return now to Gert's case, where the permissible options are:

- (a) Make a sacrifice for an altruistic reason.
- (b) Do not make the sacrifice [italicized letters are mine]. (2003, 13)

<sup>33</sup> I believe that a number of MSAs, containing *different* proportions of relaxing acts, altruistic acts, and career-furthering acts, can all be rationally permissible. I'll have more to say about this in §§ 6.5.2.



Gert denies that this is a multiple-option case, for he claims that “[s]uch pairs cannot plausibly be construed as alternate ways of pursuing the same end or acting on the same reasons” (2003, p. 13). However, he gives no argument for this claim; he must just think that it’s obvious. But it’s not at all obvious. Indeed, as I’ve shown, options *a* and *b* could be two alternative ways of pursuing the end of performing some rationally permissible MSA. Furthermore, options *a* and *b* can be construed as two alternative ways of acting on the same reason: specifically, the imperfect reason that I have to perform any set of actions that constitutes a means to my performing a rationally permissible MSA. Indeed, Gert’s options *a* and *b* could just be my options A and B above.

Of course, if securitism is to account for the basic belief, it must do more than just account for options *a* and *b* being rationally optional, it must also account for the fact that they remain rationally optional even after the strength and/or number of reasons in favor of just one of them has increased slightly. Suppose, for instance, that I’m choosing between (*a*) spending today volunteering for Oxfam and (*b*) spending today watching football on TV. Assume that both are rationally optional, for let’s suppose that it doesn’t matter which I do today so long as I spend a sufficient amount of time doing each sort of activity over time. Let’s further suppose that although Oxfam always provides each volunteer with a delicious cookie, today and *today only* it’s going to be a freshly baked one. Here, then, there’s a special reason for me to spend today as opposed to some later day volunteering for Oxfam: namely, I’ll receive a more delicious cookie if I do.<sup>34</sup> So the question arises: if, previously, there was a rational option for me to spend today either volunteering for Oxfam or watching football on TV when Oxfam wasn’t offering a freshly baked cookie, why isn’t the former now rationally required given that there is this new and additional reason for me to do so? Can securitism account for these two alternatives being rationally optional both before and after the addition of this new reason for spending today volunteering for Oxfam?

Interestingly, securitism can and can do so without appealing to satisficing, incommensurability, or anything else similarly controversial—provided, that is, that we accept the following two plausible assumptions, which I label ‘A1’ and ‘A2’:

<sup>34</sup> Whereas there is no reason to prefer receiving a cookie today to receiving one on some later day, there is a reason to prefer receiving a fresh-baked cookie today to receiving a pre-packaged one on some later day—or so we’ll assume.

- A1** If it is rationally permissible for an agent with options  $\alpha_i$  and  $\alpha_j$  to perform whichever one she presently prefers to perform (that is, whichever one she feels like performing), and if it is rationally permissible for her present preference to shift from one option to the other, then it is true to say that she has, in the ordinary sense, the rational option of performing either  $\alpha_i$  or  $\alpha_j$  even though she is, in some stricter sense, rationally required to perform the preferred option.
- A2** Some slight increase in the number and/or strength of the reasons that favor the non-preferred option is typically insufficient to outweigh the strength of the reasons that the agent has to perform the preferred option.<sup>35</sup>

To illustrate the plausibility of A1 and A2, consider the following example. Suppose that Franklin is deciding whether to order soup or salad with his entrée. Assume that the two are equal in terms of cost, taste, calories, and nutritional content, but that Franklin just happens to prefer soup to salad on this particular occasion—perhaps, he just feels like having something hot. Even if we think that, given his current preference for soup, it would be objectively irrational for him to choose the salad, we still think that he has the rational option of choosing the salad instead, for we think that he is rationally permitted to come to have the opposite preference and that, were he to come to have the opposite preference, it would be rationally permissible for him to choose the salad.<sup>36</sup> Given this, we should accept A1. And we should also accept A2 as well, for even if we were to change the example so that the salad is slightly more tasty or nutritional than the soup, we wouldn't think that this slight increase in the strength of the reasons that favor the non-preferred option would be sufficient to outweigh the strength of the reasons that Franklin has to go with his preferred option.

Given A1 and A2, securitism can account for the fact that the two options remain rationally optional even after there is some slight increase

<sup>35</sup> If no kind of desire could ever provide an agent with a reason for action, then we should take 'the preferred option' to stand for 'the option that would provide the agent with the most enjoyment given her present likes and dislikes'. For an argument that the kind of desire that's at play here can provide an agent with a reason for action, see CHANG 2004.

<sup>36</sup> This account of rational options is one that I borrow from BRATMAN 1994, pp. 330-1. See also PORTMORE 2003.

in the strength and/or number of the reasons that favor just one of them. Consider, again, the case where I'm choosing whether to spend today either volunteering for Oxfam or watching football on TV, and assume that all the rationally permissible MSAs that are, of this moment, securable by me include my spending today doing whichever one I prefer to be doing. So if, on the one hand, I prefer to watch TV today and volunteer for Oxfam on some later day, then all the rationally permissible MSAs that are presently securable by me include my watching TV today. But if, on the other hand, I prefer to volunteer for Oxfam today and watch TV (or do something else relaxing) on some later day, then all the rationally permissible MSAs that are presently securable by me include my volunteering for Oxfam today.

Now let's suppose that, in fact, I prefer to watch TV today. In that case, all the rationally permissible MSAs that are presently securable by me include my watching TV today. And even if Oxfam offers the new enticement of a freshly baked cookie, it will still be that all the rationally permissible MSAs that are presently securable by me include my watching TV today, for, as A2 states, this slight increase in the strength of the reasons that favor my volunteering for Oxfam today is insufficient to outweigh the strength of the reasons that I have to spend today doing what I prefer to be doing—namely, watching football on TV.

Of course, the reader may wonder: "In what sense is there a rational option to volunteer for Oxfam today if all the rationally permissible MSAs that are presently securable by me include my watching TV instead?" The answer is: "in the ordinary sense." As A1 states, if I'm rationally permitted to do whichever one I presently prefer to be doing, and if I'm rationally permitted to prefer either one, then it's true to say that I have the rational option to do either. That's just what it means to have a rational option. And, as I've just argued, the antecedents specified in A1 are true even on securitism. So even on securitism watching TV and volunteering for Oxfam are rationally optional both before and after the addition of this extra reason to volunteer for Oxfam today. It seems, therefore, that securitism can account for the fact that the two options can remain rationally optional (in the ordinary sense given by A1) even after there is a slight increase in the number and/or strength of the reasons that favor just one of them.

Things get a bit more interesting, however, when we up the stakes. Suppose, for instance, that Oprah Winfrey is offering, today and today only, to give to each person who volunteers a brand new car. If so, it would seem foolish of me to pass up this unique opportunity to receive a free car, even if I would, other things being equal, prefer to watch TV today and

volunteer for Oxfam on some later day. Here, then, there seems to be more than just a slight increase in the strength and/or number of the reasons that favor my volunteering for Oxfam today. Thus, whatever reason I have to spend today doing what I prefer to be doing isn't going to be strong enough to outweigh the reason that I now have to spend it volunteering for Oxfam. So securitism would seem to imply that I am now, given Winfrey's offer, rationally required to spend today volunteering for Oxfam. And this seems to be the intuitively correct result.

But, now, consider a variation on this case, one in which Bill Gates is offering, today and today only, to donate an extra \$200 to Oxfam for each person who volunteers today. And let's assume that this extra \$200 would be enough to provide measles immunizations for an additional ten children. It would seem that, here too, the new reason for volunteering today constitutes more than just a slight increase in the strength and/or number of reasons for doing so. And, thus, securitism would seem to imply that I am rationally required to volunteer for Oxfam today even if I would, other things being equal, prefer to just stay home and watch football on TV. This, I suspect, will seem counterintuitive, but let me suggest that securitism needn't have this implication, as I'll now explain.

Securitism will have this counterintuitive implication only if the additional reason that Gates's offer provides constitutes more than just a slight increase in the strength of the reasons for my spending today volunteering for Oxfam. But it seems to me that the strength of my reasons for take advantaging of such an opportunity varies greatly depending on whether or not my doing so is essential to my taking the welfare of others as one of my ultimate ends. If, on the one hand, my taking advantage of this particular opportunity is essential to my counting as someone who takes the welfare of others as one of my ultimate ends, then I will likely have most reason to complete an MSA in which I take advantage of the opportunity that Gates's offer provides. But if, on the other hand, my taking advantage of this particular opportunity isn't essential to my counting as someone who takes the welfare of others as one of my ultimate ends, then I will likely have most reason to complete whichever MSA would best further my interests. Or so I will argue.

It seems to me, as well as to others who have been similarly inspired by Kant's remarks on imperfect duties, that we are both morally and rationally required to adopt the welfare (or happiness) of others as an

ultimate end.<sup>37</sup> That is, we are required “to make the happiness of others a serious, major, continually relevant, life-shaping end” (HILL 2002, p. 206).<sup>38</sup> And this requirement in turn entails a further one: to act throughout our lives so as to exhibit a sufficient propensity to promote the happiness of others. Having such a propensity is, after all, constitutive of having adopted the happiness of others as an ultimate end. If, for instance, someone with the usual abilities and opportunities never does anything to promote the happiness of others, this would show that she has not truly adopted the happiness of others as an ultimate end (HILL 2002, p. 204).

It’s important to note, though, that having a propensity to *A* does not imply that one takes advantage of every favorable opportunity to *A*, as Robert Noggle explains:

[T]o say that Fred has a propensity to drink alcohol does not mean that Fred drinks whenever possible or even at every favorable opportunity. Fred may retain a propensity to drink even if he sometimes turns down drinks at parties that are good opportunities for drinking. Thus, one or a few failures to *A* even at good opportunities for *A*ing need not raise any doubts about a person having a propensity to *A*. However a pattern of continual failures to *A* despite favorable opportunities *would* cast doubt on the claim that a person has a propensity to *A*. (2009, p. 7)

Likewise, it seems that unless I have exhibited a pattern of failing to take advantage of good opportunities to help others, I can turn down this particular opportunity to help others by volunteering for Oxfam today and still count as having adopted the welfare of others as an ultimate end, and this is so even if this represents a particularly good opportunity for me to benefit others in that Gates is going to donate an extra \$200 if I do so. So, in this case, it seems that I would count as someone who takes the welfare of others as one of his or her ultimate ends whether I spend today volunteering for Oxfam or watching football on TV. My volunteering for Oxfam today simply isn’t essential to my counting as someone who takes the welfare of others as one of his or her ultimate ends.

And it seems to me that, in those instances in which I’ll count as having taken the welfare of others as an ultimate end either way, the primary

<sup>37</sup> I borrow the term ‘an ultimate end’ from Robert Noggle, who uses it to refer to “an intrinsic end that is a fundamental and indispensable part of the agent’s life” (2009, p. 8). An agent’s ultimate ends may include such things as the happiness of others, the well-being of her loved ones, and the perfection of herself.

<sup>38</sup> I think that we’re also required to adopt other things as ultimate ends—things such as our own perfection and the happiness and perfection of our children.

determiner of whether I should perform an MSA that includes my spending today volunteering for Oxfam or one that includes my spending today watching football on TV is which activity would best promote my self-interest, which in turn is likely determined by which I happen to prefer to be doing today.<sup>39</sup> Here, we come close to what James Griffin calls a discontinuity in value. Take two values *A* and *B*. These values are *discontinuous* if “so long as we have enough of *B* any amount of *A* outranks any further amount of *B*” (1986, p. 85). So long as we restrict ourselves to MSAs in which I count as taking the welfare of others as an ultimate end, then, other things being equal, the rationally permissible MSAs are the ones that are best in terms of my self-interest.<sup>40</sup>

So, despite initial appearances, it seems that the additional reason that Gates’s offer provides for my volunteering for Oxfam today amounts to no more than a slight increase in the strength of the reasons to do so—at least, this is so assuming that I have not regularly failed to take advantage of good opportunities to promote the welfare of others and that I plan to perform a set of actions that will ensure that I perform an MSA in which I take advantage of such opportunities sufficiently often.<sup>41</sup> And thus it is both this fact and A2 that explains why my spending today watching TV remains rationally permissible even after Gates provides me with an additional reason to instead volunteer for Oxfam.

Of course, it may be that were Gates to offer to donate enough extra money, it would, then, cease to make sense for me to spend today watching

<sup>39</sup> This doesn’t mean that that it will be objectively rational for me to continue to put off helping others until my doing so is no longer consistent with my having adopted the welfare of others as an ultimate end. In some instances, it will be in my self-interest to help others sooner rather than later, as where, for instance, I prefer volunteering for Oxfam now and relaxing later to relaxing now and volunteering for Oxfam later.

<sup>40</sup> The talk of ‘enough beneficence’ can be misleading. The idea is not that there is some cap on how much beneficence can be required of us such that once we’ve exceeded this cap we can no longer be required to do anything more for the sake of others. This is implausible, for it seems that no matter how much we have already done for the sake of others, we could still be required to do more, as where, for instance, we happen upon a child drowning in a shallow pond, whom we can save at very little cost to ourselves. So although we do need to be sufficiently beneficent to count as having adopted the welfare of others as an ultimate end, the fact that the duty of beneficence consists in the duty to make the welfare of others “a serious, major, continually relevant, life-shaping end” (HILL 2002, p. 206) means that this duty is never over and done with—see NOGGLE 2009, esp. p. 12.

<sup>41</sup> If you have exhibited a pattern of failing to take advantage of good opportunities to help others, then there will come a point at which you can no longer continue to do so and still count as having adopted the welfare of others as an ultimate end. At this point, taking advantage of this opportunity is, given your past neglect, rationally required.

TV, for it may be that certain opportunities are so exceptionally good that one could not count as having adopted the welfare of others as an ultimate end and yet pass them up.<sup>42</sup> But the chance to ensure that Oxfam receives an additional \$200 doesn't seem to be this sort of "golden opportunity," as there are typically plenty of other such opportunities to ensure that Oxfam receives an additional \$200.<sup>43</sup>

These are, at any rate, my views on the matter. If readers have other views about how to assess the rational permissibility of MSAs, they should apply those views when evaluating securitism. The point has been only to show that securitism needn't imply that, in the case in which Gates offers to contribute an extra \$200 for each volunteer, I am rationally required to volunteer for Oxfam today even if I would, other things being equal, prefer to just stay home and watch football on TV. Of course, some readers may not agree with some of my assumptions, but, regardless, they should think that securitism has counterintuitive implications only if they think either (1) that there are instances in which an agent is rationally permitted to perform a certain act even though she is rationally prohibited from performing any MSA in which she performs that act or (2) that there are instances in which an agent is rationally prohibited from performing a certain act even though she is rationally permitted to perform some MSA in which she performs that act. So, for instance, one should think that securitism has counterintuitive implications in the Gates Case only if one thinks both that I'm rationally permitted to stay home and watch TV and that there is no rationally permissible MSA in which I do so.

### **§6.5 Securitism's suppositions and implications**

In this section, I explore securitism further. Doing so will help us to better understand the theory as well as to stave off some potential

<sup>42</sup> Also, if the example were changed so that I could ensure that Oxfam receives an extra \$200 by sacrificing merely the fulfillment of some very trivial preference (such as, the slight preference I have to write out addresses in print as opposed to cursive), I might, then, be required to make this small sacrifice (thereby sacrificing some very small portion of my self-interest) even if my doing so isn't essential to my leading a full and meaningful life. As I say above, it seems to me that even if one has enough beneficence in one's life to count as leading a life that's full and meaningful, a life that's moderately better in terms of beneficence may outrank a life that's only the slightest bit better in terms of one's self-interest.

<sup>43</sup> Again there is no cap on the amount of beneficence that can be required of us. Thus no matter how beneficent one has been in the past, there may be certain "golden opportunities" that everyone is required to take advantage of—see NOGGLE 2009, p. 12.

misconceptions about the view. I start off by considering whether securitism must make any implausible suppositions to account for the basic belief. To account for the basic belief, securitists must assume that what matters most is that agents perform MSAs that achieve a reasonable balance between various worthy ends and that it often matters very little, if at all, in which order these ends are pursued. Some might worry, then, that securitists must implausibly suppose that the best way to balance the various competing ends will be the same for each of us and that, for each of us, there will be only one way to optimally balance them. I show that this is not the case. I also show that although securitism presupposes that MSAs can be obligatory, it does not presuppose that agents can form the intention to perform each and every act in which large and complex MSAs consists. Lastly, I turn to securitism's implications with regard to end-of-life decisions, procrastination, and Warren Quinn's self-torturer puzzle. I argue that securitism has plausible implications in each case.

*6.5.1 Does securitism implausibly suppose that the best way to balance the various competing factors in our lives will be the same for each of us?* We all have to balance various competing factors in our lives: e.g., career, family, entertainment, personal projects, and the welfare of others. But it would be a mistake to suppose that the best way to balance these competing factors is going to be the same for each of us. It's not as if the best MSA for each of us is going to contain the same proportions of various worthy pursuits, as where, say, 40% of the acts in such an MSA would be focused on furthering one's career, another 40% would be focused on our family and friends, and the remaining 20% would be focused on helping others. As I see it, there are no universally ideal proportions.

What's important is that each of us leads a life that's full and meaningful, but there are many different paths to such a life. Living a full and meaningful life requires that each of us to adopt certain ultimate ends—the welfare of others being one of them. But the fact that we should each have this as our end doesn't entail that we should each dedicate the same percentage of our resources to helping others. Indeed, someone with abundant resources and limited familial obligations will need to plan on performing MSAs in which a much higher percentage of her actions are focused on helping others than, say, someone else with limited resources and a disabled child to care for.

So we're not required to dedicate the same percentage of our resources to any given end. Nor are we required to adopt the same specific set of ends. Even if there are certain broadly defined ends that we should each adopt in order to live full and meaningful lives, these ends will, in practice,



yield different derivative ends for different people. Take the end of developing one's talents. For one person, it will mean adopting the development of one's musical gifts as an end. For another, it will mean adopting the development of one's philosophical talents as an end. So, different people will need to take different paths to full and meaningful lives. Moreover, of those MSAs that one could perform so as to put one on the path to leading a full and meaningful life, some will be more choice-worthy than others given one's individual psychological traits. Consider, for instance, that although we all need enjoyment in our lives, we enjoy different things. Whereas some find skydiving exhilarating, others find it terrifying.

In sum, the best way to balance various competing factors in our lives is going to vary from person to person, for a lot is going to depend on each person's individual circumstances and psychological traits. In virtue of this, which MSAs, with which sorts of proportions of various sorts of activities, will be rationally permissible will vary from one individual to another.

*6.5.2 Does securitism implausibly suppose that, for each of us, there is only one way to optimally balance the various competing factors in our lives?* As noted above, how any individual should optimally balance the various completing factor in her life is going to depend a great deal on her own individual circumstances and psychological traits. But even given a particular set of circumstances and psychological traits, there can be more than one way to optimally balance two or more competing factors. For instance, trading \$30,000 for a Honda Accord might, for a given individual, be just as good as trading \$70,000 for a Jaguar XJ. Each of these two trades may strike, what is for her, an optimal tradeoff between savings and luxury. Likewise, there may be more than one way for a person to strike an optimal balance between, say, altruism and self-interest. Imagine, for instance, a young doctor who must choose to work either for The Mount Sinai Medical Center in New York or for Doctors without Borders in Africa. The former offers the chance to live in greater material comfort and to employ more sophisticated medical techniques, whereas the latter trades less of these goods for the greater opportunity to do good. Both, then, involve a certain tradeoff between altruism and self-interest, but both tradeoffs may be just as good. So, on securitism, it may be that, for each person, there is a range of different tradeoffs among various competing goods that all count as optimal.

*6.5.3 Does securitism implausibly suppose that agents can form the intention to perform each and every act in which some very large and complex MSA*

*consists?* Securitism presupposes that MSAs can be permissible, impermissible, and even obligatory. Yet, in most instances, it would be impossible for a human agent to intend to perform each and every act in which such MSAs consist. This is because many of these MSAs include literally billions of acts performed over multiple decades. Given the limits of human comprehension, no human could ever form the intention to perform each and every act contained within such a large and complex MSA.

A human agent could, however, form the intention to achieve an end that requires her to perform such a very large and complex set of actions. To illustrate, consider that, when I was younger, I formed the intention to earn a PhD in philosophy and that this was an end that I could achieve only by performing a very large and complex set of actions. At the time, I had only the vaguest notion of all that I would need to do to succeed. So even though I did not form the intention to perform some highly-specified set of actions culminating in my receiving a PhD in philosophy, I did form the vague intention to do pretty much whatever it takes (within reason) to earn a PhD in philosophy. And this intention, we'll suppose, is what led me down the path that culminated in my having performed the very large and complex set of actions needed to earn a PhD in philosophy. These sorts of vague intentions are not at all beyond human limits, and this is all that the proponent of securitism needs. On securitism, I'm obligated to perform some MSA, because there is some rather limited intention (say, the intention to pursue a PhD in philosophy) that will, given counterfactual determinism, determine that I perform the obligatory MSA.

*6.5.4 Is it problematic that securitism implies that as we approach the end of our lives our rational options narrow considerably?* Suppose that I'm deciding whether to spend this Sunday at the beach with my family or at home alone watching football on TV. On securitism, both would typically be rationally optional, for there would typically be both a rationally permissible MSA that's presently securable by me in which I spend today watching TV and some later day doing something with my family and a rationally permissible MSA that's presently securable by me in which I spend today doing something with my family and some later day watching TV (or doing something else similarly relaxing). What's important is that I take advantage of both opportunities to relax and opportunities to spend time with my family and that I do each sufficiently often, but it doesn't generally matter which I do on any particular occasion.

But now suppose this is my last day of life. In that case, I have no future opportunities beyond today. And so the choice is no longer between

doing one of these activities today and the other on some later day, but is instead between doing one of them today and the other never again. Given this, I tend to think that my current preferences are not going to be decisive in determining which option is rationally permissible, even though they might have been decisive had the choice been between which to do today and which to do on some later day. I'm inclined to think that if my choice is between spending my last day with my family and spending it alone watching TV, I ought spend it with my family even if, not knowing that my death is imminent, I happen to prefer to spend the day alone watching TV.

Far from being objectionable, this implication seems to be a merit of the view. After all, I can't see how any fully informed and well-meaning advisor would recommend that I spend my last day alone watching TV. Surely, there is something better I could be doing on my last day! So I think that securitism is right to imply that as we approach the end of our lives our rational options narrow considerably.

But now consider a different sort of case. Suppose that I'm on my death bed dictating my last will and testament. And let's assume that I've already provided for my family's continual financial security, as I have taken out a rather substantial life-insurance policy with them as the beneficiaries. Some might claim that, having provided for my family's financial security, I am rationally permitted to do almost anything with my estate that would not be harmful, leaving it to any individual or organization or even using it to purchase some extravagant funeral for myself—one in which, say, my body is blasted out into space (see COCHRANE 2009).

I'm not so sure. First, let's take the case where I have so far exhibited throughout my life a sufficient concern for others, having clearly adopted the welfare of others as one of my ultimate ends. Since having such a concern for others doesn't necessitate taking advantage of every good opportunity to help others, I could pass up this (my last) opportunity to help others and still count as having led a full and meaningful life. Given this, it seems that it might be rationally permissible for me to use the money for my own benefit by, say, creating an open-access, digital archive of my work. Of course, it might also be permissible for me to donate my entire estate to charity. Which of these two options would in the end be best would depend, I think, on which would best serve my self-interests, which in turn would depend on whether I happen to care more about helping others to an even greater extent than I already have or more about ensuring that my work is made widely available. I don't, however, think that it would be rationally permissible for me to leave my estate to some random well-to-do stranger or to spend it on some extravagant funeral in

which my body is blasted out into space. Given my particular cares and concerns, either action would seem positively foolish.

Now, consider the case in which I haven't lived a life in which I've exhibited having adopted the welfare of others as one of my ultimate ends. Unfortunately, nothing I can do now will change this. I was required to have adopted the happiness of others as "a serious, major, continually relevant, life-shaping end" (HILL 2002, p. 206), and I didn't. The best I can do now is to atone for this moral failure. So the question becomes, "Which course of action is more choice-worthy: the one in which I atone for my previous selfishness by leaving most, if not all, of my estate to some charitable organization or the one in which I continue to act for my own benefit, using the money to set up a digital archive of my work?" I think that the former is more choice-worthy. And if this is right, then (maximizing) securitism implies that it would be rationally impermissible for me to use the money for my own benefit.<sup>44</sup>

Of course, all this is based on the assumption that no MSA that is presently securable by me in which I use the money for my own benefit instead of donating it to charity would be rationally permissible. If the reader thinks otherwise, then she should think that securitism has different implications. For instance, if it turns out that I'm wrong and that the MSA in which I use the money to set up the digital archive of my work is in fact rationally permissible, then securitism will imply that using my estate to do this is rationally permissible. Again, the reader should think that securitism is unacceptable only if her views on what it is, as of present, rationally permissible (or impermissible) to do is incompatible (or compatible) with her views about which of the presently securable MSAs it would be rationally permissible to perform.

*6.5.5 What does securitism imply about procrastination?* Securitism has plausible implications with regard to the rational permissibility of procrastinating. To illustrate, suppose that there are two things I really need to get done today: exercising and finishing some grading. Let's suppose that I usually exercise in the morning, because I find that if I put it off until the afternoon, I often lack the will power to exercise. But let's suppose that I really prefer to finish my grading this morning and put off exercising until this afternoon, which is when the weather is supposed to

<sup>44</sup> Maximizing securitism holds that a given MSA is objectively rationally permissible if and only if there is not more reason to perform some alternative MSA. We don't have to adopt maximizing securitism, but it is interesting to see how even maximizing securitism can account for the basic belief.

be ideal for a run. Whether it is permissible for me to put off exercising until this afternoon depends, according to securitism, on whether there is some plan, resolution, or other intention that I could form now that would ensure that I finish grading this morning and exercise this afternoon. If there is, then that's what I should plan on doing. If there isn't, and if the only way to ensure that I both exercise and finish grading today is to plan on exercising this morning and finish grading this afternoon, then that's what I ought to do.

Securitism, thus, plausibly supposes that whether it is objectively rationally permissible to put something off for later depends on whether one can effectively intend to do so. If one's present intention will be effective in ensuring that one succeeds in doing it later, then it's permissible to put it off. But if there is no present intention that would ensure that one succeeds in doing it later, then one should not put it off and should instead do it now.

*6.5.6 What does securitism imply about cases such as Warren Quinn's case of the self-torturer? Consider Warren Quinn's famous puzzle of the self-torturer:*

Suppose there is a medical device that enables doctors to apply electric current to the body in increments so tiny that the patient cannot feel them. The device has 1001 settings: 0 (off) and 1 . . . 1000. Suppose someone (call him the self-torturer) agrees to have the device, in some conveniently portable form, attached to him in return for the following conditions: The device is initially set at 0. At the start of each week he is allowed a period of free experimentation in which he may try out and compare different settings, after which the dial is returned to its previous position. At any other time, he has only two options -- to stay put or to advance the dial one setting. But he may advance only one step each week, and he may *never* retreat. *At each advance he gets \$10,000.*

Since the self-torturer cannot feel any difference in comfort between adjacent settings, he appears to have a clear and repeatable reason to increase the voltage each week. The trouble is that there *are* noticeable differences in comfort between settings that are sufficiently far apart. Indeed, if he keeps advancing, he can see that he will eventually reach settings that will be so painful that he would then gladly relinquish his fortune and return to 0.

The self-torturer is not alone in his predicament. Most of us are like him in one way or another. We like to eat but also care about our appearance. Just one more bite will give us pleasure and won't make us look fatter; but very many bites will. And there may be similar connections between puffs of pleasant smoking and lung cancer, or between pleasurable moments of idleness and wasted lives. (1990, p. 79)

As Quinn notes, although the self-torturer's situation is highly unusual, most of us face similar predicaments on a regular basis. It is important,

then, for a theory of objective rationality to offer plausible suggestions as to what one should do when facing such predicaments. Securitism does. To see this, let's assume that the self-torturer will, after the 1,000 weeks are over, regret having advanced the dial as many times as he did if and only if he advances the dial no more than  $n$  times. Given Quinn's stipulations, it's clear that  $n < 1,000$ . Now suppose that the self-torturer faces the decision of whether to advance the dial a step during some week short of  $n$  weeks into it. According to securitism, the self-torturer should advance a step this week if and only if there is some intention to  $\varphi$  such that, if the self-torturer were to intend now to  $\varphi$ , he would end up advancing no more than  $n$  times altogether. Suppose, then, that it's the first week and the setting is at 0. Suppose that the self-torturer has played around with various settings and determined that the best trade off between pain and money lies with advancing the dial exactly  $n$  steps. Suppose, then, that he resolves to advance the dial the next  $n$  weeks and to stop after the  $n^{\text{th}}$  week. There are two possibilities. Perhaps, his resolution will hold. Perhaps, he will advance  $n$  weeks, and come the following week he will not reconsider his intention to advance only  $n$  steps and will stand firm at  $n$  steps. If this is what would happen, then the self-torturer should advance a step each of the next  $n$  weeks.

However, there's another possibility. After advancing the first  $n$  weeks, the self-torturer might be tempted to reconsider his earlier intention to stop after the  $n^{\text{th}}$  week. He might, then, try advancing the dial a step the following week and find that there is no noticeable difference. Finding that there is no noticeable difference and wanting an additional \$10,000, perhaps he will reconsider his earlier intention and decide to advance another step beyond  $n$  steps. Perhaps, he continues to do this. In the end, he ends up in so much pain that he would gladly return all the money to return to the 0 setting. If this is what would happen, then securitism implies that the self-torturer should not advance a step this week. Thus, securitism plausibly implies that whether the self-torturer should advance a step this week depends on whether his resolution to advance no more than  $n$  steps will hold firm even if he advances this week. If it will, he should advance. If it will not, he shouldn't. And this all seems quite plausible.



In this chapter, I've argued that we can account for the basic belief without appealing to any controversial assumptions such as the assumption that, in

most choice situations, the relevant reasons are incommensurable or have two separable dimensions of strength (justifying strength and requiring strength). I've argued that we can account for the basic belief by appealing to only that which all sides agree on: that there are multiple-option cases and that in such cases the relevant imperfect reasons provide equal reason for pursuing each of the equally attractive means to achieving the relevant end. I've argued that, ultimately, what agents have most reason to do is to perform rationally permissible MSAs. And the reason an agent has to perform some rationally permissible MSA is an imperfect reason, for it equally supports all of the various other particular acts that each constitutes a means to performing some rationally permissible MSA. Since, in most choice situations, many of the available act alternatives constitute a way of acting on this imperfect reason, we find that the relevant reasons permit performing any of these various act alternatives. And this, I've argued, is how we can account for the basic belief.

In the process of making these arguments, I have developed and defended a new theory of objective rationality, which I call securitism. I've argued that this theory is promising not only because it enables us to account for the basic belief, but also because it provides plausible verdicts as to whether to procrastinate and as to what to do when facing a predicament such as the one that Quinn's self-torturer faces. Moreover, securitism allows us to meet Kagan's objection while holding onto the account of agent-centered options that I defended in the previous chapter.

In the next chapter, I apply the lessons learned in this chapter—in particular, that 'obligatory' implies 'securable', that we need to be able to assess not only the deontic statuses of minimal acts but also that of MSAs and of everything in between, and that permissions and obligations must be time-indexed—to moral theory, arguing that they have important implications for how we should best formulate utilitarianism as well as consequentialism in general.

## CHAPTER 7

---

# Commonsense Consequentialism

- 7.1 The best version of utilitarianism: commonsense utilitarianism
- 7.2 Securitist consequentialism and the argument for it
- 7.3 Commonsense consequentialism and how it compares with traditional act-consequentialism
- 7.4 What has been shown and what remains to be shown

IN THIS CHAPTER, I return to the theory with which I began the book: maximizing act-utilitarianism. Applying the lessons learned over the last six chapters, I suggest a series of incremental improvements on the theory, eventually arriving at what I take to be the most plausible version of utilitarianism: *commonsense utilitarianism*. I then show how, by abstracting from this theory, we can arrive at what I take to be the most plausible version of consequentialism: *commonsense consequentialism*. This theory is, I believe, quite promising. I show just how promising it is via a comparison with traditional act-consequentialism. Lastly, I end the chapter and the book by summarizing what has been shown and explaining what remains to be shown.

### §7.1 The best version of utilitarianism: commonsense utilitarianism

§§7.1.1 *Maximizing act-utilitarianism*: One of the simplest and most traditional consequentialist theories is *maximizing act-utilitarianism* (MAU). Its criterion of rightness is as follows:

- MAU<sub>cr</sub>** It is objectively morally permissible for S to perform an act, *x*, if and only if there is no alternative act that's personally possible for S that would produce more aggregate utility than *x* would.

MAU is unsatisfactory for a number of reasons, not the least of which is that it conflicts with moral rationalism given certain plausible assumptions about what agents have decisive reason to do—see Chapter 2. But it also fails for a couple of technical reasons.



First, it fails to index permissions to times despite the fact that it makes permissibility a function of what's personally possible for the agent, which is something that changes over time. To illustrate the problem, suppose that  $x$  is, as of  $t_1$ , the best alternative that's personally possible for me. MAU implies, then, that I'm obligated to perform  $x$  and that my performing  $y$ , an inferior alternative, is impermissible. But suppose that  $t_3$  rolls around and things have changed such that  $x$  is no longer personally possible for me and that  $y$  is now (as of  $t_3$ ) the best alternative that's personally possible for me. Surely, it can't be that I'm still obligated to perform  $x$ . After all, 'obligatory' implies 'personally possible' and  $x$  is no longer personally possible for me. So is  $x$  obligatory or not? And is  $y$  permissible or impermissible? It seems impossible to answer such questions without indexing permissions (and hence also obligations) to times, which is something MAU fails to do.

Second, MAU fails to provide criteria for assessing the deontic statuses of various sets of actions despite the fact that we deliberate about them and form intentions to perform them. After all, MAU is mostly silent with regard to the deontic statuses of sets of actions, for the variable ' $x$ ' ranges over only individual acts (types and tokens). Of course, MAU does render a verdict with regard to certain sets of actions: specifically, those that constitute some larger, individual act. Take, for instance, the act of firing a semi-automatic handgun. This compound act is comprised of a sequence of more basic actions: inserting the magazine, releasing the slide, taking aim, and squeezing the trigger. Insofar as this sequence of actions constitutes the larger, compound act of firing a semi-automatic handgun, MAU renders a verdict as to its deontic status. But not all sets of actions make up some larger, compound action. For instance, the set of actions that I intend to perform today—drop off my daughter at school in the morning, return home to pack, leave the house by 11 a.m., stop to get gas, head down to Tucson, grab some lunch along the way, pick up my friend Dave at Tucson International Airport, and then head over to the Westward Look Resort for the workshop that we'll be attending—does not constitute any individual action. Thus, MAU is silent as to its deontic status. Yet this is a set of actions about which I've deliberated. Moreover, it's a set of actions that I intend to perform. Surely, then, an adequate moral theory should tell me whether it is a set of actions that it would be permissible for me to perform.

§§7.1.2 *Revised maximizing act-utilitarianism*: These two technical problems can be easily fixed. We need only to adopt what I call *revised maximizing act-utilitarianism* and its criterion of rightness:

**RMAU<sub>cr</sub>** It is, as of  $t_i$ , objectively morally permissible for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if there is no alternative set of actions that is, as of  $t_i$ , personally possible for S that would produce more aggregate utility than  $\alpha_j$  would.

RMAU does index permissions to times, and it does allow us to assess the deontic statuses of all possible sets of actions. Nevertheless, it is not without its own problems.

One problem is that it violates the following plausible principle of deontic logic, which holds that *permissibility distributes over conjunction* (hence, 'PDC'):

**PDC**  $P(S, t_i, [x_1, x_2, \dots, \& x_n]) \rightarrow [P(S, t_i, x_1), P(S, t_i, x_2), \dots, \& P(S, t_i, x_n)]$ ,

where, if you'll recall, ' $P(S, t_i, \varphi)$ ' stands for 'S is, as of  $t_i$ , permitted to  $\varphi$ '. To see this, consider the following revised version of the procrastination example from Chapter 6:

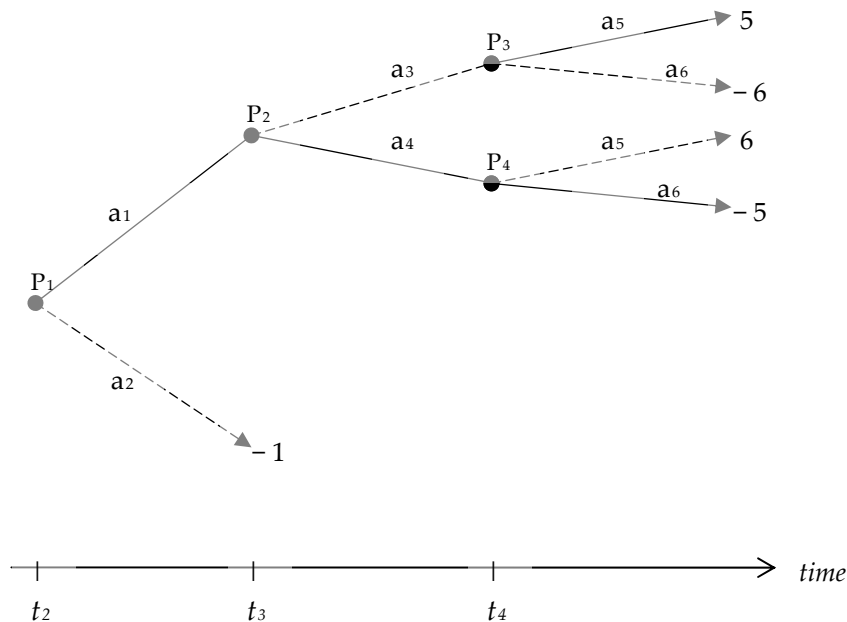
*Professor Procrastinate II:* Professor Procrastinate receives, at  $t_1$ , an invitation to review a book. He is the best person to do the review and has the time to do it. If he both accepts the invitation at  $t_2$  and then teaches a seminar on the book at  $t_3$ , he will then write the review at  $t_4$ , which is what would be best both for him and for others. However, he is, as a matter of fact, going to teach, at  $t_3$ , a course on logic as opposed to a seminar on the book—the result being that he's going to become so engrossed at  $t_4$  in redesigning his logic course that he'll put off writing the review indefinitely. Assume, then, that although Professor Procrastinate's accepting the invitation at  $t_2$ , teaching a seminar on the book at  $t_3$ , and writing the review at  $t_4$  is, as of  $t_1$ , personally possible for him, it is not, as of  $t_1$ , securable by him. For there is, as of  $t_1$ , nothing that he can do or intend to do that will secure his performance of this course of action. Indeed, no matter how resolute he is, at  $t_1$ , in his intention to teach a seminar on the book at  $t_3$  and write the review at  $t_4$ , he will instead end up being overcome both by the temptation to teach logic at  $t_3$  and by the temptation to redesign his logic course at  $t_4$ . And, thus, he will end up perpetually procrastinating, failing to ever write the review. It would be better, then, if he were to decline the invitation at  $t_2$ . For it would be better to do this than to accept the

invitation at  $t_2$  and end up never writing the review, which is precisely what would happen were he to accept the invitation.

The situation is depicted graphically in Figure 7.1. The points on the graph labeled 'P<sub>1</sub>', 'P<sub>2</sub>', and 'P<sub>3</sub>' represent choice points. The branches stemming from each choice point represent the set of mutually exclusive and jointly exhaustive act alternatives that are, as of that point, securable by Professor Procrastinate—assume, if necessary, that God will cause Professor Procrastinate to cease to exist should he intend to do anything other than the acts depicted in the graph and that, therefore, the branches on the graph represent his *only* options. Thus, at P<sub>1</sub>, he must choose to perform either a<sub>1</sub> or a<sub>2</sub>, but not both.

**Figure 7.1**

- |  |  |
|--|--|
| a <sub>1</sub> = accept the invitation       | a <sub>4</sub> = teach logic               |
| a <sub>2</sub> = decline the invitation      | a <sub>5</sub> = write the book review     |
| a <sub>3</sub> = teach a seminar on the book | a <sub>6</sub> = redesign the logic course |



Solid lines represent acts that Professor Procrastinate would freely choose to perform were he to reach the relevant choice point. Dashed lines

represent acts that Professor Procrastinate could, but would not, freely choose to perform were he to reach the relevant choice point. Thus, if Professor Procrastinate were to reach  $P_3$ , he would, as the solid line represents, freely choose to perform  $a_5$ , not  $a_6$ . But he could perform  $a_6$  in that, were he to intend at  $t_4$  to perform  $a_6$ , that's what he would do.

Time is represented by the line at the bottom. So, as one can see from the graph,  $a_1$ 's time-interval extends from  $t_2$  to  $t_3$ , and  $a_6$ 's time-interval extends from  $t_3$  to  $t_4$ .

Lastly, the numerals to the right of each arrow represent the utility of the corresponding compound acts. Thus, Professor Procrastinate's performing  $a_1$  would produce  $-5$  utiles, for Professor Procrastinate would end up performing the compound act,  $a_1$ -and- $a_4$ -and- $a_6$ , if he were to perform  $a_1$ . Likewise, we can see from the graph that his performing  $a_1$ -and- $a_3$  would produce  $5$  utiles, for he would end up performing the compound act,  $a_1$ -and- $a_3$ -and- $a_5$ , if he were to perform  $a_1$ -and- $a_3$ .

According to RMAU, it is, as of  $t_1$ , objectively morally impermissible for Professor Procrastinate to perform  $a_1$ , as there is an alternative—viz.,  $a_2$ —that would produce more utility than  $a_1$  would. Professor Procrastinate's performing  $a_1$  would produce  $-5$  utiles, whereas his performing  $a_2$  would produce  $-1$  utile. Now, although RMAU implies that Professor Procrastinate is not, as of  $t_1$ , objectively morally permitted to perform  $a_1$ , RMAU implies that he is, as of  $t_1$ , objectively morally permitted (indeed required) to perform  $a_1$ -and- $a_3$ , as the only alternative to this compound act is  $a_1$ -and- $a_4$ , which is suboptimal. Professor Procrastinate's performing  $a_1$ -and- $a_4$  would produce  $-5$  utiles, whereas his performing  $a_1$ -and- $a_3$  would produce  $6$  utiles.<sup>1</sup> Since RMAU implies that Professor Procrastinate is, as of  $t_1$ , objectively morally permitted to perform  $a_1$ -and- $a_3$  but is not, as of  $t_1$ , objectively morally permitted to perform  $a_1$ , it's clear that RMAU violates PDC. And violating PDC is, as I explained in Chapter 6, a high price to pay for accepting any theory.

What's even worse, though, is that RMAU implies that agents should pander to their own present imperfection. That is, it permits agents to perform worse alternatives simply because they don't have at present the intentions that they ought to have. To see this, consider yet another version of the procrastination example from Chapter 6:

<sup>1</sup> Professor Procrastinate's performing  $a_1$ -and- $a_4$  is the only alternative to his performing  $a_1$ -and- $a_3$ , for alternatives must be both agent-identical and time-identical. Two sets of actions are *agent-identical* if and only if they have the same agent. And two sets of actions are *time-identical* if and only if they persist over the exact same interval (or intervals) of time.

*Professor Procrastinate III:* Professor Procrastinate receives, at  $t_1$ , an invitation to review a book. He is the best person to do the review and has the time to do it. If he both accepts the invitation at  $t_2$  and then teaches a seminar on the book at  $t_3$ , he will write the review at  $t_4$ , which is what would be best both for him and for others. However, as a matter of fact, he's going to teach logic at  $t_3$ , and this will result in his becoming so engrossed in redesigning his logic course at  $t_4$  that he'll put off writing the review indefinitely. Now, the only reason that he's going to end up teaching logic at  $t_3$  is because that's what he intends at  $t_1$  to do. If he were instead to intend at  $t_1$  to teach a seminar on the book, that's what he would do. Nevertheless, given his actual intention at  $t_1$  to teach logic, it's true to say that what he would do were he to accept the invitation is to teach logic at  $t_3$ , redesign his logic course at  $t_4$ , and never get around to writing the review.

The situation is depicted graphically in Figure 7.2. The faint dash-dotted line between  $P_1$  and  $P_2$  represents the fact that  $a_1 \& i_2$  isn't actually an option that Professor Procrastinate can choose, at will, to perform, for  $i_2$  is an intention, not an action. Intentions, unlike actions, are not under an agent's volitional control. An agent can't choose, at will, what to intend any more than she can choose, at will, what to believe.<sup>2</sup> Professor Procrastinate's choice, then, is between performing  $a_1$  and performing  $a_2$ , and not between performing  $a_1$  with this or that intention.

Nevertheless, intentions, like beliefs, are attitudes for which there can be reasons. Moreover, intentions are judgment-sensitive attitudes, meaning that they are responsive to our judgments about reasons. So, it's fair to say that Professor Procrastinate ought to intend at  $t_1$  to teach a seminar on the book at  $t_3$ , for he has decisive reason to do so. And, thus, Professor Procrastinate's failure to have this intention doesn't excuse him from being obligated to perform what is, as of  $t_1$ , the best course of action presently securable by him: accepting the invitation at  $t_2$ , teaching a seminar on the book at  $t_3$ , and writing the review at  $t_4$ .<sup>3</sup> To secure his performance of this

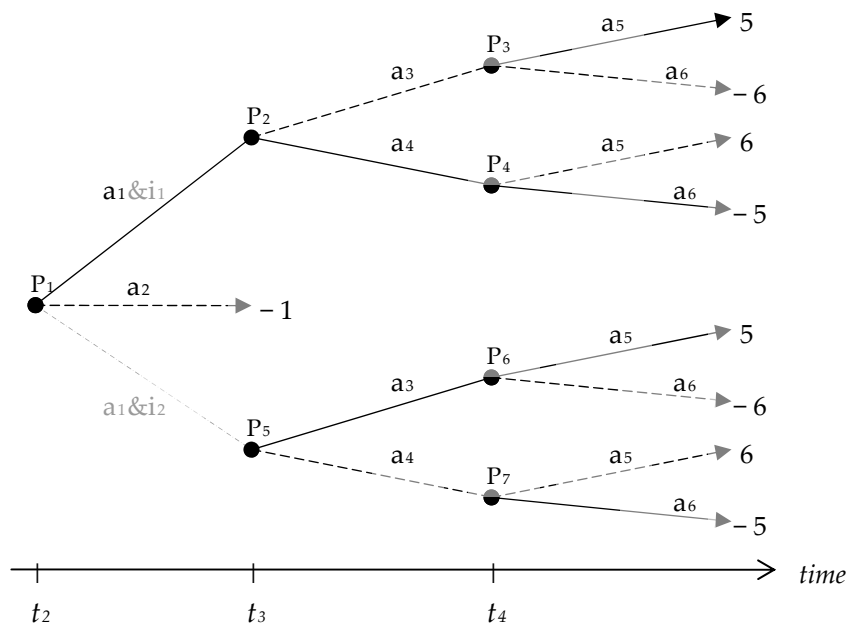
<sup>2</sup> Of course, agents can choose to perform acts that will cause them to have certain beliefs. Suppose, for instance, I could pay someone to hypnotize me into believing that P. This, then, would be a case where I chose to perform an act that caused me to believe that P. But, in this case, it's the physical act of my paying this person, not the mental "act" of my believing that P, that I chose, at will, to perform.

<sup>3</sup> I'm assuming that his writing the review would be sufficiently beneficial both for him and for others that there is decisive reason for him to write the review.

(the best) course of action, Professor Procrastinate need only to intend at  $t_1$  to perform this course of action, which is something that he ought to intend at  $t_1$  to do. The problem, then, is only that Professor Procrastinate intends at  $t_1$  to teach logic, when he ought to intend at  $t_1$  to teach a seminar on the book.

**Figure 7.2**

- |   |                                     |
|---|-------------------------------------|
| $i_1$ = intend to teach logic                 | $a_3$ = teach a seminar on the book |
| $i_2$ = intend to teach a seminar on the book | $a_4$ = teach logic                 |
| $a_1$ = accept the invitation                 | $a_5$ = write the book review       |
| $a_2$ = decline the invitation                | $a_6$ = redesign the logic course   |



Implausibly, RMAU implies that Professor Procrastinate's failure to intend at  $t_1$  to teach a seminar on the book excuses him from being obligated to perform what is, as of  $t_1$ , the best course of action presently securable by him. This is because RMAU is a version of the following implausible view.

*Actualism:* It is, as to  $t_i$ , objectively (morally or rationally) permissible for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if the set of actions that S would actually perform were S to perform  $\alpha_j$  is

no worse than the set of actions that S would actually perform were S to perform any alternative set of actions that is, as to  $t_i$ , personally possible for S.

The problem with actualism is that it takes as given certain future actions that an agent will perform only because she does not at present have the intentions that she currently ought to have. But an agent shouldn't take as given those future actions that are under her present volitional control. For instance, Professor Procrastinate shouldn't take as given that he will be teaching logic at  $t_3$  when whether he's going to teach logic at  $t_3$  is presently up to him and whether he presently intends to do so or not.

In other words, when Professor Procrastinate deliberates about whether to accept the invitation, he shouldn't reason to the conclusion that he ought to decline the invitation on the assumption that he would never get around to writing the review if he were to accept the invitation. For whether or not this assumption is correct depends on what he presently intends to teach at  $t_3$ , and, in deliberating about what to do, he should not hold fixed future actions that are under his present volitional control (Carlson 1999, p. 265). So rather than think that he should decline the invitation because he's going to teach logic at  $t_3$  and thereby become engrossed at  $t_4$  with revising his logic course instead of writing the book review, he should think that he ought to accept the invitation and decide now to teach a seminar on the book so that he won't be distracted at  $t_4$  from writing the review.

§§7.1.3 *Maximizing securitist utilitarianism*: So although RMAU avoids some of the technical problems that plagued MAU, it is clearly not the best version of utilitarianism. It has counterintuitive implications in cases such as *Professor Procrastinate II* and *Professor Procrastinate III*. A view that avoids such counterintuitive implications as well as the technical problems that plagued MAU is *maximizing securitist utilitarianism*:

**MSU** (a) It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as

of  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by S whose outcome contains more aggregate utility than  $MSA_i$ 's outcome does.

Like RMAU, MSU rightly implies that, in *Professor Procrastinate II*, Professor Procrastinate is not, as of  $t_1$ , objectively morally permitted to perform  $a_1$ . But, unlike RMAU, MSU doesn't counterintuitively imply that PDC is false, for, unlike RMAU, MSU doesn't further imply that Professor Procrastinate is, as of  $t_1$ , objectively morally permitted to perform  $a_1$ -and- $a_3$ . In *Professor Procrastinate II*, none of the objectively morally permissible maximal sets of actions that are, as of  $t_1$ , securable by Professor Procrastinate includes his performing  $a_1$ -and- $a_3$  for the simple reason that his performing  $a_1$ -and- $a_3$  is not, as of  $t_1$ , securable by him. So MSU is superior to RMAU in that it doesn't violate PDC.

MSU is also superior RMAU in that it doesn't imply that agents should pander to their own present imperfection. In *Professor Procrastinate III*, MSU, unlike RMAU, implies that Professor Procrastinate is not, as of  $t_1$ , objectively morally permitted to perform  $a_2$ . The MSAs that are, as of  $t_1$ , securable by Professor Procrastinate are: (1)  $a_1$ -and- $a_4$ -and- $a_6$  with -5 utiles, (2)  $a_2$ -and- $a_4$ -and- $a_6$  with -1 utiles, and (3)  $a_1$ -and- $a_3$ -and- $a_5$  with 5 utiles. According to MSU, then, the only objectively morally permissible MSA is  $a_1$ -and- $a_3$ -and- $a_5$ , which produces the most utility of the three. Thus, Professor Procrastinate is not, as of  $t_1$ , objectively morally permitted to perform  $a_2$ , since there is no objectively morally permissible MSA that is, as of  $t_1$ , securable by Professor Procrastinate which includes his performing  $a_2$ .

So MSU is clearly superior to its actualist counterpart—viz., RMAU. But besides actualism and securitism, there is also the following view:

*Possibilism:* It is, as to  $t_i$ , objectively (morally or rationally) permissible for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if the best maximal set of actions that is, as to  $t_i$ , personally possible for S includes S's performing  $\alpha_j$ .

To determine whether MSU is the best version of utilitarianism, we need to know not only whether it's superior to its actualist counterpart (RMAU),



but also whether it is superior to its possibilist counterpart—viz., *maximizing possibilist utilitarianism*:

**MPU** (a) It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , personally possible for S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , personally possible for S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , personally possible for S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , personally possible for S whose outcome contains more aggregate utility than  $MSA_i$ 's outcome does.

MPU is inferior to MSU. The problem with MPU, and with possibilism in general, is that it is unrealistic. As Holly Goldman notes, “[t]here seems little point in prescribing an act which puts the agent in a position to do great things if the same act also puts him in a position to do something disastrous, and he would choose the latter rather than the former” (1976, p. 468). Consider again *Professor Procrastinate II*. MPU implies that Professor Procrastinate is, as of  $t_1$ , objectively morally required to accept the invitation (that is, to perform  $a_1$ ), because this would put him in the position to perform the best subsequent course of action—viz.,  $a_4$ -and- $a_5$ . But why should Professor Procrastinate put himself in the position to perform the best subsequent course of action—viz.,  $a_4$ -and- $a_5$ —if this also puts him the position to perform the worst subsequent course of action—viz.,  $a_4$ -and- $a_6$ —and, as a matter of fact, he is going to choose to perform the latter rather than the former? I could see the point in MPU's prescribing that Professor Procrastinate performs  $a_1$  now if there was something that he could do, or intend to do, now that would change the fact that he would follow up by performing  $a_4$ -and- $a_6$  instead of  $a_4$ -and- $a_5$ . But, in *Professor Procrastinate II*, there is nothing that he can do, or intend to do, now to change this fact. And, in that case, it seems pointless for MPU to prescribe that he performs  $a_1$ . After all, what's the point of prescribing that he performs  $a_1$  so that he will end up in the position to perform the best

subsequent course of action when, in fact, there is nothing he can do, or intend to do, now that will prevent him from following up with the worst subsequent course of action?

Another problem with MPU, and with possibilism in general, is that it implausibly implies that there is the following sort of asymmetry: how some future agent would behave in response to S's  $\varphi$ -ing is relevant to whether S should  $\varphi$  if that future agent is someone other than S, but not if that future agent is S. But why should it matter whether that future agent is S or someone else if regardless there is nothing that S can do, or intend to do, now that will change how this future agent is going to behave in response to S's  $\varphi$ -ing? If, in neither case, does S have present deliberative control over what that future agent will do, why should the two cases be treated differently?<sup>4</sup>

Of course, only in those instances in which that future agent is S is it possible for S to have present deliberative control over what that future agent is going to do. And sometimes S does have such control, as where, say, S will  $\psi$  at some future time,  $t_3$ , if and only if, and because, S intends at  $t_1$  to  $\psi$  at  $t_3$ . In such instances, S has present deliberative control over whether or not she will  $\psi$  at  $t_3$ . But even if S does sometimes have present deliberative control over whether or not she will  $\psi$  at  $t_3$ , it's important to realize that this is not always the case. For, sometimes, no matter what S intends now, at  $t_1$ , to do, S is going to  $\psi$  at  $t_3$ . What's relevant, then, is not whether the future agent is S or someone else but whether that future agent's behavior is under S's present deliberative control. If it's not, then, contrary to possibilism, whether S should  $\varphi$  depends on what that future agent would do were S to  $\varphi$ .

Thus, possibilism, unlike securitism, fails to distinguish between those cases in which S has present deliberative control over what she does in the future and those cases in which she does not. It focuses instead on the irrelevant distinction between whether the future agent is S or someone else. Thus, there is, according to possibilism, no relevant distinction to be made between Professor Procrastinate's situation in *Professor Procrastinate II* and his situation in *Professor Procrastinate III*. In both cases, it is true to say that Professor Procrastinate's accepting the invitation to review the book at  $t_2$  and writing the review at  $t_4$  is, as of  $t_1$ , personally possible for him. And, on possibilism, that's all that matters. The fact that, in *Professor Procrastinate II*, Professor Procrastinate would not write the review at  $t_4$  if

<sup>4</sup> A number of philosophers have made this point. See, for instance, GOLDMAN 1976 (p. 469) and WOODARD 2009 (pp. 219-220).

he were to accept the invitation at  $t_2$  is, according to possibilism, completely irrelevant even though there is absolutely nothing that he can, as of  $t_1$ , do to alter this fact. But surely this unalterable fact is relevant. Professor Procrastinate should, as of  $t_1$ , treat his future procrastination as unalterable if there is, as of  $t_1$ , nothing that he can do, or intend to do, at  $t_1$  to ensure that he doesn't procrastinate at  $t_4$ . It's only in *Professor Procrastinate III* that Professor Procrastinate should treat his future procrastination as alterable. For it is only in this case that there is, as of  $t_1$ , some way for Professor Procrastinate to ensure that he doesn't procrastinate and writes the review at  $t_4$ —specifically, by intending at  $t_1$  to teach a seminar on the book at  $t_3$ .

Possibilism also ignores the way we in fact reason about practical matters. We often reason that we ought not to  $\varphi$  because  $\varphi$ -ing would put us in a position in which we're very likely to give into the temptation and do something even worse than not  $\varphi$ -ing (GOLDMAN 1976, pp. 469-470). Thus, in *Professor Procrastinate II*, Professor Procrastinate should, knowing his tendency to procrastinate, reason that he ought to decline the invitation. And we often reason that we ought to plan now on  $\psi$ -ing in the future since our  $\psi$ -ing in the future will prevent us from being tempted to do the wrong thing. Thus, in *Professor Procrastinate III*, Professor Procrastinate should reason that he ought to intend now to teach a seminar on the book at  $t_3$  so that he won't be tempted at  $t_4$  to redesign his logic course as opposed to write the review. It seems, then, that possibilist theories such as MPU fail to capture the way that we in fact reason about practical matters. Thus, MSU seems superior to MPU not only in that it rightly distinguishes between those cases in which one's future behavior is under one's present deliberative control and those cases in which it is not, but also in that it better captures the way we in fact reason about practical matters.

So we've seen that MSU is superior to both its actualist and possibilist counterparts: that is, to RMAU and MPU, respectively. Nevertheless, I don't think that MSU is, in its present formulation, the best version of utilitarianism. For, as it's currently formulated, it cannot account for the basic features of commonsense morality: agent-favoring options, agent-sacrificing options, special obligations, agent-centered restrictions, supererogatory acts, and the self-other asymmetry. We should, then, apply the lessons that we learned in Chapter 4 and reformulate MSU so as to accommodate these features. When we do, we get what I call *commonsense utilitarianism*.

§§7.1.4 *Commonsense utilitarianism*: Commonsense utilitarianism is that version of utilitarianism that accommodates more of our commonsense

moral intuitions than any other version of utilitarianism does. The following is an initial stab at formulating this view:

CSU<sub>1</sub> (a) It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions, MSA <sub>$i$</sub> , is objectively morally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S. (c) MSA <sub>$i$</sub> 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by S whose outcome contains both more constraint-adjusted utility and more comprehensively-adjusted utility than MSA <sub>$i$</sub> 's outcome does, where the constraint-adjusted utility of an outcome is just the sum of the utility for others, adjusted by multiplying any disutility (or loss of utility) resulting from S's infringements of an agent-centered constraint by 500, and where the comprehensively-adjusted utility of an outcome is just its constraint-adjusted utility added to the product of S's utility times ten.

This is only an initial stab at formulating commonsense utilitarianism. Commonsense utilitarianism will undoubtedly need to be more nuanced than this. But even if CSU<sub>1</sub> is rather simplistic, it's good enough for our present purposes. CSU<sub>1</sub> illustrates how a utilitarian theory can accommodate the basic features of commonsense morality: agent-favoring options, agent-sacrificing options, special obligations, agent-centered restrictions, supererogatory acts, and the self-other asymmetry. In this respect it's like SU from Chapter 4. But it is superior to SU in that it is securitist, indexes permissions to times, and provides criteria for assessing everything from a minimal act to a maximal set of actions.

To illustrate how CSU<sub>1</sub> can accommodate the basic features of commonsense morality, consider the scenario depicted in Table 7.1. In this scenario, there are four mutually exclusive and jointly exhaustive maximal

sets of actions that are, as of  $t_0$ , securable by an agent named Edouard:  $MSA_1$ ,  $MSA_2$ ,  $MSA_3$ , and  $MSA_4$ .

**Table 7.1**<sup>5</sup>

Let  $U_i(\varphi)$  = the disutility (or loss of utility) that will accrue to others as a result of  $S$ 's infringing on any agent-centered constraint by performing  $\varphi$ , let  $U_o(\varphi)$  = the utility that will accrue to others, excluding  $U_i(\varphi)$ , as a result of  $S$ 's performing  $\varphi$ , let  $U_{con}(\varphi) = U_o(\varphi) + [500 \times U_i(\varphi)]$ , let  $U_s(\varphi)$  = the utility that will accrue to  $S$  as a result of  $S$ 's performing  $\varphi$ , and let  $U_{com}(\varphi) = U_{con}(\varphi) + [10 \times U_s(\varphi)]$ .

$MSA_i$	$U_o(\varphi)$	$U_i(\varphi)$	$U_{con}(\varphi)$	$U_s(\varphi)$	$U_{com}(\varphi)$	deontic status
$MSA_1$	10	0	<b>10</b>	7	<b>80</b>	merely perm.
$MSA_2$	24	-1	<b>-476</b>	50	<b>24</b>	impermissible
$MSA_3$	15	0	<b>15</b>	1	<b>25</b>	supererogatory
$MSA_4$	20	0	<b>20</b>	-1	<b>10</b>	supererogatory

$CSU_1$  accommodates agent-centered options in that it permits agents to give their own utility anywhere from no weight at all up to ten times the weight of any other. Thus, Edouard can permissibly choose to do what is better for himself and perform  $MSA_1$  even though his performing  $MSA_4$  would be both better for others and better in terms of the amount of impersonal good produces.<sup>6</sup> Edouard has, then, the agent-favoring option of performing  $MSA_1$  instead of  $MSA_4$ . Furthermore, by allowing agents to give their own utility no weight,  $CSU_1$  accommodates agent-sacrificing options, which permit agents to forgo their own greater good for the sake of providing others with some lesser net benefit. For instance, Edouard can permissibly choose to perform  $MSA_3$  as opposed to  $MSA_1$ , thereby providing others with a net benefit of five utiles at a cost of six utiles to himself.<sup>7</sup>

<sup>5</sup> I've emphasized the fourth and sixth columns with bold-faced type since these two represent the two auxiliary rankings of which permissibility is, according to  $CSU_1$ , a function.

<sup>6</sup>  $U_o(MSA_1)$  is 10, whereas  $U_o(MSA_4)$  is 20. And the total utility produced by Edouard's performing  $MSA_1$  is 17, whereas the total utility produced by Edouard's performing  $MSA_4$  is 19.

<sup>7</sup> The benefit to others is  $U_o(MSA_3) - U_o(MSA_1)$ , which equals 5. The cost to herself is  $U_s(MSA_3) - U_s(MSA_1)$ , which equals -6.

Thus, CSU<sub>1</sub> accommodates what's known as *the self-other asymmetry*: the asymmetry found in commonsense morality between what agents are permitted to do to themselves and what agents are permitted to do to others. It seems that although an agent is not permitted to be indifferent toward the welfare of others, she is permitted to be indifferent toward her own welfare. Edouard can, for instance, sacrifice his hand to save Fay's thumb even though it would be impermissible for him to sacrifice Gustav's hand to save Fay's thumb (even if Gustav consents to his doing so). On commonsense morality, then, agents are permitted to sacrifice their own greater good for the sake of providing others with some lesser net benefit, but they are not permitted to sacrifice someone else's greater good for the sake of providing others with some lesser net benefit—hence, the asymmetry between self and others. Of course, agents are not always permitted to sacrifice their own welfare. It would, for instance, be wrong, as CSU<sub>1</sub> implies, for Edouard to choose to perform MSA<sub>2</sub> as opposed to MSA<sub>1</sub>. Such a choice would not only be worse for Edouard, but worse for others as well.<sup>8</sup>

CSU<sub>1</sub> can also accommodate agent-centered constraints, which include both special obligations and agent-centered restrictions. According to CSU<sub>1</sub>, there will, for instance, be a non-absolute restriction against committing murder. Agents would, other things being equal, be prohibited from committing murder even so as to prevent up to 500 others from each committing some comparable murder. And, according to CSU<sub>1</sub>, there will be a non-absolute constraint against failing to fulfill the special obligation that parents have to nurture their own children. Parents would, other things being equal, be prohibited from failing to nurture their own children even so as to prevent up to 500 others from each comparably failing to nurture their children. Clearly, then, CSU<sub>1</sub> can accommodate both agent-centered restrictions and special obligations.

CSU<sub>1</sub> can even accommodate the notion of going above and beyond the call of duty—or, at least, it can if the following is a plausible account of what makes a maximal set of actions supererogatory:

**Super-Max** It is, as of  $t_i$ , supererogatory for S to perform a maximal set of actions, MSA <sub>$j$</sub> , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, there is some alternative maximal set of actions, MSA <sub>$i$</sub> , also beginning at  $t_j$  ( $t_i \leq t_j$ ) such that: (a) it

<sup>8</sup> Whereas  $U_s(\text{MSA}_1)$  is 7,  $U_s(\text{MSA}_2)$  is only 2. And whereas  $U_o(\text{MSA}_1)$  is 10,  $U_o(\text{MSA}_2)$  is only 4.

is, as of  $t_i$ , both objectively morally permissible for S to perform  $MSA_j$  and objectively morally permissible for S to perform  $MSA_i$  and (b) S has more moral reason to perform  $MSA_j$  than to perform  $MSA_i$ .

Certain maximal sets of actions will, on  $CSU_1$ , meet conditions *a* and *b*. To see this, note that  $CSU_1$  has two auxiliary rankings: ( $AR_1$ ) a ranking of outcomes in terms of their constraint-adjusted utility, and ( $AR_2$ ) a ranking of outcomes in terms of their comprehensively-adjusted utility. As I see it,  $AR_1$  is a ranking of outcomes in terms of how much moral reason the agent has to bring them about. It specifies that the more constraint-adjusted utility an outcome contains the more moral reason there is to want it to obtain and, thus, to act so as to bring it about. And  $AR_2$  of  $CSU_1$  is a ranking of outcomes in terms of how much reason, all things considered, the agent has to bring them about. It specifies that the more comprehensively-adjusted utility an outcome contains the more reason there is, all things considered, to want it to obtain and, thus, to act so as to bring it about.

If this is right, then Edouard has more moral reason to perform  $MSA_3$  than he has to perform  $MSA_1$ , a permissible alternative, for  $MSA_3$  produces more constraint-adjusted utility than  $MSA_1$  does. And the same holds for  $MSA_4$ .<sup>9</sup> Thus  $MSA_3$  and  $MSA_4$  are both supererogatory on  $CSU_1$ . What's more,  $CSU_1$  allows for a range of supererogatory alternatives, where some supererogatory alternatives go further beyond the minimum required than others do.<sup>10</sup> For instance,  $MSA_4$  involves going further beyond the minimum that morality requires than  $MSA_3$  does. In this scenario, Edouard is required to produce at least 10 units of constraint-adjusted utility, and  $MSA_4$  produces 10 additional units beyond this minimum, whereas  $MSA_3$  produces only 5 additional units beyond this minimum.

§§7.1.5 *Supererogatory versus superperfecterogatory*: Unfortunately, Super-Max doesn't provide criteria for determining whether some non-maximal set of actions is supererogatory. So we can't yet assess whether  $CSU_1$  accommodates the intuitive idea that various non-maximal sets of

<sup>9</sup>  $U_s(a_3)$  and  $U_s(A_4)$  are 15 and 20, respectively, whereas  $U_s(a_1)$  is only 10.

<sup>10</sup> On commonsense morality, there is often a range of supererogatory acts, where some supererogatory acts go further beyond what's required than others. For instance, it seems that not only is my choosing to spend my Saturdays helping the poor supererogatory, but so is my choosing to spend both my Saturdays and my Sundays helping the poor. Although both choices would seem to be supererogatory, the latter seems to go further beyond what's required of me than the former does.

acts can be supererogatory. Perhaps, though, the problem can be easily solved by adopting the following principle, which holds that *supererogatoriness distributes over conjunction*:<sup>11</sup>

$$\text{SUPDC } \text{SUP}(S, t_i, [\alpha_1, \alpha_2, \dots, \& \alpha_n]) \rightarrow [\text{SUP}(S, t_i, \alpha_1), \text{SUP}(S, t_i, \alpha_2), \dots, \& \text{SUP}(S, t_i, \alpha_n)],$$

where ‘ $\text{SUP}(S, t_i, \alpha_i)$ ’ stands for ‘ $S$ ’s performing  $\alpha_i$  is, as of  $t_i$ , supererogatory’.

In fact, this won’t do, for SupDC is clearly false. SupDC is false in virtue of the fact that there are imperfect duties, which can be fulfilled only by performing certain sets of actions over time. Thus, it’s possible for a set of actions to go beyond the call of some imperfect duty even though none of the acts within that set go beyond the call of duty themselves. Take, for instance, the (imperfect) duty of beneficence. In general, it doesn’t require that you take advantage of any particular opportunity to help others, but it does require that you take advantage of such opportunities with sufficient regularity—at least, assuming that you have certain typical abilities and opportunities. So, for instance, even if it would clearly go beyond the call of duty for Kimo to donate \$200 to charity each and every month given that this is all the disposable income that he has, his donating \$200 on any particular occasion would not in and of itself go beyond the call duty. If, in the past thirty years, the only thing that Kimo has ever done to help others is to donate \$200 on January 1, 2000, then this act of charity would not itself count as going above and beyond the call of duty. Indeed, given that this is all that he’s ever done, he has surely fallen well short of what’s minimally required by the (imperfect) duty of beneficence. The fact that he has done so little despite both the means and opportunity to do considerably more demonstrates that he has not made “the happiness of others a serious, major, continually relevant, life-shaping end” (HILL 2002, p. 206), which is what the duty of beneficence requires.

So, given the nature of the duty of beneficence, that it is an imperfect duty, a duty to adopt the happiness of others as “a serious, major, continually relevant, life-shaping end,” it follows that an agent can go beyond the call of the duty of beneficence only by performing a temporally-extended set of actions that goes beyond what is minimally required to count as having adopted the welfare of others as “a serious, major, continually relevant, life-shaping end.” But although such a set of

<sup>11</sup> Paul McNamara has noted this fact—see 1996a, p. 432.



actions would be supererogatory, it doesn't follow that any of the individual acts within that set goes beyond the call of duty. Thus, SupDC is false. And this means that what I had, in earlier chapters, been calling a 'supererogatory act' is more accurately called a 'superperfecterogatory act', i.e., an act that goes beyond the call of only what perfect duty requires.<sup>12</sup> More precisely, a *superperfecterogatory act* is defined as follows:

**SUPERF** It is, as of  $t_i$ , superperfecterogatory for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, there is some alternative set of actions,  $\alpha_i$ , also beginning at  $t_j$  ( $t_i \leq t_j$ ) such that: (a) it is, as of  $t_i$ , both objectively morally permissible for S to perform  $\alpha_j$  and objectively morally permissible for S to perform  $\alpha_i$  and (b) S has more moral reason to perform  $\alpha_j$  than to perform  $\alpha_i$ .

Kimo's donating \$200 to charity on January 1, 2000, was superperfecterogatory, for he was not morally required to do so. He was, we'll suppose, morally permitted to have instead spent that \$200 on himself.<sup>13</sup> But although this act was superperfecterogatory, it was not supererogatory. Kimo did not thereby go beyond the call of duty, for he did not thereby exceed what the duty of beneficence requires. So, in contrast to superperfecterogatory acts, we can define a *supererogatory set of actions*—a set that goes beyond the call of both perfect and imperfect duty—as follows:

**SUPER** It is, as of  $t_i$ , supererogatory for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, both of the following hold: (1) in performing  $\alpha_j$ , S does not thereby minimally or partially fulfill some positive duty, and (2) there is some alternative set of

<sup>12</sup> I borrow this term and the distinction from SINNOTT-ARMSTRONG 2005, p. 204.

<sup>13</sup> Not all permissible acts are superperfecterogatory. To illustrate, suppose that Lorenzo and Melissa are both drowning and that I can save only one of the two, as they are so far apart that there is no way to reach both of them in time. Let's further suppose that I'm morally obligated to save one of the two. And let's assume that there is no more moral reason to save the one than the other. Finally, assume that, after tossing a coin to decide whom to save, I save Melissa. Saving Melissa was permissible, but it wasn't superperfecterogatory, for, although there was another alternative that I was permitted to do instead (viz., to save Lorenzo), I had no more moral reason to save Melissa than to save Lorenzo.

actions,  $\alpha_i$ , also beginning at  $t_j$  ( $t_i \leq t_j$ ) such that: (a) it is, as of  $t_i$ , both objectively morally permissible for S to perform  $\alpha_j$  and objectively morally permissible for S to perform  $\alpha_i$  and (b) S has more moral reason to perform  $\alpha_j$  than to perform  $\alpha_i$ .<sup>14</sup>

There are a number of aspects to this definition that are in need of clarification. First,  $\alpha_i$  and  $\alpha_j$  are alternatives if and only if they have the same agent and extend for the same intervals of time—that is, if and only if they are both agent-identical and time-identical. Second, in  $\varphi$ -ing, S does thereby only minimally fulfill some positive duty, D, if and only if S's  $\varphi$ -ing fulfills D and there is no more moral reason for S to  $\varphi$  than to do anything else that would fulfill D. Third, in  $\varphi$ -ing, S does thereby partially fulfill some positive duty, D, if and only if S's  $\varphi$ -ing is a proper subset of some set of actions by which S minimally fulfills D. Fourth, a positive duty is one that can be fulfilled only by performing some set of actions. This contrasts with a negative duty, which can be fulfilled only by refraining from performing certain actions.

To see both how SUPER yields intuitively plausible verdicts in a range of cases and how CSU<sub>1</sub> can accommodate these plausible verdicts and thereby accommodate various supererogatory non-maximal sets of actions, it will be helpful to work through a few examples. Assume that none of the morally relevant details have been left out of the descriptions of these cases. Assume that certain alternatives have been left out only when they are vastly inferior to the others and, thus, morally irrelevant. And assume, as noted above, that the more constraint-adjusted utility an outcome contains the more moral reason there is for the agent to want it to obtain and, thus, to act so as to bring it about. Lastly, assume that the imperfect duty of beneficence is a kind of agent-centered constraint: specifically, a kind of special obligation that one may not permissibly infringe upon even so as to prevent more numerous others from comparably infringing upon it. Assume, then, that it is impermissible to fail to adopt the happiness of others as “a serious, major, continually relevant, life-shaping end” even if doing so will prevent more numerous others from doing the same. Here, then, are the cases.

*Case 1 (The Generous Donation):* Suppose that over the next twelve months I donate \$120 per month to Oxfam even though my circumstances are such that I could minimally fulfill my duty of beneficence by donating,

<sup>14</sup> This definition is inspired in part by MELLEMA 1991.

on average, \$100 per month over my entire adult life.<sup>15</sup> Further assume that I don't minimally or partially fulfill any positive duty by donating an additional \$20 a month over these next twelve months. Assume, then, that I haven't, for instance, promised anyone that I would donate an additional \$20 a month. And assume that I will always give, and have always given, at least \$100 a month to charity. Thus, I'm not, for instance, obligated to give more in these coming months so as to make up for giving less than \$100 in past months. Lastly, assume that the relevant alternatives and their associated utilities are as depicted in Table 7.2.

**Table 7.2**

Let  $\alpha_1$  = my donating \$90 per month over the next twelve months, let  $\alpha_2$  = my donating \$100 per month over the next twelve months, and let  $\alpha_3$  = my donating \$120 per month over the next twelve months.

$\alpha_i$	$U_o(\varphi)$	$U_i(\varphi)$	$U_{con}(\varphi)$	$U_s(\varphi)$	$U_{com}(\varphi)$	deontic status
$\alpha_1$	90	-10	-4,910	40	-4,870	impermissible
$\alpha_2$	100	0	100	20	300	merely permissible
$\alpha_3$	120	0	120	0	120	supererogatory

In this case, SUPER rightly implies that my donating an additional \$20 a month over the next twelve months is supererogatory. And CSU<sub>1</sub> accommodates this verdict. My performing  $\alpha_3$  is supererogatory, for CSU<sub>1</sub> implies that  $\alpha_2$  and  $\alpha_3$  are both permissible and that there is more moral reason for me to perform  $\alpha_3$  than to perform  $\alpha_2$ , as  $\alpha_3$  produces more constraint-adjusted utility.

*Case 2 (The Minimal Donation):* Assume that everything is as it is in *The Generous Donation* except that in this case I donate an average of only \$100 a month over the next twelve months. Again, the relevant alternatives and their associated utilities are as depicted in Table 7.2.

In this case, SUPER rightly implies that my donating an average of only \$100 a month over the next twelve months is merely permissible, and not supererogatory. And CSU<sub>1</sub> accommodates this verdict, for although CSU<sub>1</sub> implies that  $\alpha_2$  is permissible, it also implies that there is no more moral

<sup>15</sup> Strictly speaking, the duty of beneficence requires only that agents adopt the welfare of others as an ultimate end, but, from this duty, it follows that those with certain abilities and opportunities have a duty to act in whatever way is constitutive of their having adopted the welfare of others as an ultimate end.

reason for me to perform it than to perform any other permissible alternative. After all,  $\alpha_2$  produces no more than the minimum amount of constraint-adjusted utility that I'm required to produce.

*Case 3 (The Burning Building):* Suppose that, at grave risk of bodily injury, I go into a burning building to rescue the two children who are trapped inside. Assume that I have not made any promise that would obligate me to do this. And assume that I fulfill my duty to promote the welfare of others by, among other things, frequently donating significant portions of my disposable income to worthy charities. Lastly, assume that the relevant alternatives and their associated utilities are as depicted in Table 7.3.

**Table 7.3**

Let  $\alpha_1$  = my calling 911 and waiting outside for the fire department to arrive, and let  $\alpha_2$  = my going into the burning building to rescue the two children trapped inside.

$\alpha_i$	$U_o(\varphi)$	$U_i(\varphi)$	$U_{con}(\varphi)$	$U_s(\varphi)$	$U_{com}(\varphi)$	deontic status
$\alpha_1$	0	0	0	0	0	merely permissible
$\alpha_2$	2,000	0	2,000	-300	-1,000	supererogatory

In this case, SUPER rightly implies that my going into the burning building to rescue the two children trapped inside is supererogatory, for, in rescuing these two children, I do not minimally or partially fulfill any positive duty. The duty of beneficence doesn't require that I ever perform such heroic deeds. And CSU<sub>1</sub> accommodates this plausible verdict. My performing  $\alpha_2$  is supererogatory, for CSU<sub>1</sub> implies that  $\alpha_1$  and  $\alpha_2$  are both permissible and that there is more moral reason for me to perform  $\alpha_2$  than to perform  $\alpha_1$ , as  $\alpha_2$  produced more constraint-adjusted utility than  $\alpha_1$  would have.

*Case 4 (The Loan):* Suppose that I owe my friend \$1,000, which is the amount that he loaned me exactly one year ago. Since that time, our fortunes have reversed. He has hit upon some difficult financial times, and I have unexpectedly inherited a fortune from my recently deceased uncle. Given our changes in fortune, I promised my wife that I would repay his loan immediately and with \$200 in interest, and this despite the fact that our original agreement didn't call for the payment of any interest and didn't call for repayment until two year's time. Regardless, I keep my

promise to my wife and immediately pay my friend \$1,200. Assume that the relevant alternatives and their associated utilities are as depicted in Table 7.4.

**Table 7.4**

Let  $\alpha_1$  = my paying my friend \$1,000 a year from now, and let  $\alpha_2$  = my immediately paying my friend \$1,200.

$\alpha_i$	$U_o(\varphi)$	$U_i(\varphi)$	$U_{con}(\varphi)$	$U_s(\varphi)$	$U_{com}(\varphi)$	deontic status
$\alpha_1$	1,000	-10	-4,000	200	-2,000	impermissible
$\alpha_2$	1,200	0	1,200	0	1,200	obligatory

In this case, SUPER rightly implies that my immediately paying my friend \$1,200 is obligatory, and not supererogatory. In immediately paying back my friend's loan with interest, I more than minimally fulfill my duty to repay his loan, but I only minimally fulfill my duty to keep my promise to my wife. And, thus, I did nothing supererogatory in immediately paying my friend \$1,200.<sup>16</sup> CSU<sub>1</sub> accommodates this verdict, for CSU<sub>1</sub> implies that the only relevant alternative, viz.,  $\alpha_1$ , is impermissible, making  $\alpha_2$  obligatory.

*Case 5 (The Mahatma):* Consider the choice that Gandhi faced early in life when he had to decide whether to return to India to practice law or to extend his stay in South Africa to fight the newly introduced Franchise Bill, which aimed at disenfranchising Indians. He chose the latter, but let's suppose that, had he instead returned to India, he would have established a very successful law practice, and he and his family would have prospered, leading very pleasant and comfortable lives. Now, the political life that he chose for himself was a difficult one, which included beatings, starvation, imprisonment, emotional turmoil, and eventually his assassination. But, of course, he achieved great things and did a lot of good for both India and the world. Had he instead chosen to lead the life of a successful lawyer, he would have achieved much less, but lived a longer and significantly more pleasant life. Assume that the relevant alternatives and their associated utilities are as depicted in Table 7.5.

**Table 7.5**

<sup>16</sup> This is, of course, consistent with its being supererogatory for me to promise my wife to repay him immediately and with interest.

Let  $\alpha_1$  = Gandhi's leading a comfortable life as a lawyer, and let  $\alpha_2$  = Gandhi's leading a difficult life as a political activist.

$\alpha_i$	$U_o(\varphi)$	$U_i(\varphi)$	$U_{con}(\varphi)$	$U_s(\varphi)$	$U_{com}(\varphi)$	deontic status
$\alpha_1$	100	0	<b>100</b>	600	<b>6,100</b>	permissible
$\alpha_2$	5,000	0	<b>5,000</b>	100	<b>6,000</b>	supererogatory

Clearly, Gandhi's choice to lead the difficult, political life that he did (that is, to perform  $\alpha_2$ ) was supererogatory, which is what SUPER implies. And CSU<sub>1</sub> accommodates this verdict. Gandhi's performing  $\alpha_2$  is supererogatory, for CSU<sub>1</sub> implies that  $\alpha_1$  and  $\alpha_2$  were both permissible options and that there was more moral reason for Gandhi to perform  $\alpha_2$  than to perform  $\alpha_1$ , as  $\alpha_2$  produced more constraint-adjusted utility than  $\alpha_1$  would have.

*Case 6 (The Daughters):* Suppose that I promised my friend Omar that I would help one of his two daughters—that is, either Paloma or Rene—with their homework. And assume that there is actually more moral reason for me to help Paloma, for she would benefit more from my help. Yet let's suppose that there is also more non-moral reason for me to help Paloma, as it is no more difficult for me to help Paloma than it is for me to help Rene and that I would, in fact, prefer to help Paloma in that I enjoy her company more. Assume that the relevant alternatives and their associated utilities are as depicted in Table 7.6.

**Table 7.6**

Let  $\alpha_1$  = my helping neither Paloma nor Rene, let  $\alpha_2$  = my helping Rene; and let  $\alpha_3$  = my helping Paloma.

$\alpha_i$	$U_o(\varphi)$	$U_i(\varphi)$	$U_{con}(\varphi)$	$U_s(\varphi)$	$U_{com}(\varphi)$	deontic status
$\alpha_1$	0	-5	<b>-2,500</b>	0	<b>0</b>	impermissible
$\alpha_2$	5	0	<b>5</b>	1	<b>15</b>	impermissible
$\alpha_3$	10	0	<b>10</b>	2	<b>30</b>	obligatory

SUPER implies that my helping Paloma is obligatory, not supererogatory. Although I have more moral reason to help Paloma than to help Rene, helping Paloma doesn't constitute a supererogatory act, for, as CSU<sub>1</sub> implies, my helping Rene is impermissible and, thus, my helping

Paloma is obligatory. Helping Rene is impermissible on CSU<sub>1</sub>, because there is an alternative (viz.,  $\alpha_3$ ) that I have both more moral reason and more reason, all things considered, to perform. I have more moral reason to help Paloma because she would benefit more from my help than Rene would. And I have more reason, all things considered, to help Paloma, because this would be better not only for Paloma but also for me given that I prefer her company.

It may seem that my helping Paloma should be optional, not obligatory.<sup>17</sup> Although I admit that it can seem this way, I'm not so sure. It is, after all, plausible to suppose that, other things being equal, it is generally wrong to sacrifice one's own greater good if this will only make things worse for others overall. And this is precisely what I would be doing if I helped Rene instead of Paloma. My helping Rene is less good for others overall, since Paloma would benefit more from my help. And it would be less good for me as well, since I would enjoy helping Paloma more than I would enjoy helping Rene. So, in helping Rene, I would be sacrificing my own greater good only to bring about less good for others. And, in general, such action seems wrong. We might think, though, that in this specific case what's at stake in my helping Paloma rather than Rene is insufficient to generate a moral requirement to help her. In essence, the claim would be that the moral reason that I have to help Paloma instead of Rene has no moral requiring strength. I'm not sure whether this is the right thing to say about this particular case, but I certainly want to allow that some moral reasons have no moral requiring strength.

Here's my thinking. There are both requiring and non-requiring moral reasons for desiring an outcome and that what an agent has requiring reason to *do* is a function of only what she has requiring reason to *want*. We might think, for instance, that although you have a moral reason to prefer, other things being equal, the outcome in which an acquaintance, named Lee, is slightly cheered by your complimenting him on his stylish shoes, this reason is not such that you would be guilty of any moral infraction if you didn't prefer this outcome to an otherwise equal outcome in which he does not receive your compliment and is not thereby cheered. Although minimally decent people will, in general, prefer that others are better off, they needn't be concerned (at least, not always) with such trivial benefits to others. This contrasts with cases in which there clearly is a requiring reason to prefer one outcome to another. Take, for instance, the preference for

<sup>17</sup> Both Michael Ferry (2009) and Jussi Suikkanen (2009a) have raised this objection against the view that I offered PORTMORE 2003.

children not starving. To lack this preference would be a significant failing. Lacking this preference indicates a serious flaw in one's moral character, and, insofar as one is responsible for failing to develop a good moral character, one could be blameworthy for lacking a preference for children not starving.<sup>18</sup> Here, then, there seems to be a requiring reason for preferring that children not starve.

If we accept that there is such a distinction between requiring and non-requiring moral reasons for desiring, then we'll need to abandon CSU<sub>1</sub>. The problem with CSU<sub>1</sub> is that it doesn't permit us to draw a distinction between those moral reasons with and those moral reasons without moral requiring strength. And that's not CSU<sub>1</sub>'s only problem. Another problem with CSU<sub>1</sub> is that it implausibly assumes that an agent's reasons for and against desiring an outcome is always a function of what people's utilities are in that outcome. To the contrary, it seems that sometimes agents have a reason to desire that an outcome not obtain, not because it's an outcome in which anyone's utility has been diminished, but because it's an outcome in which a constraint has been violated. For instance, it seems that an agent has a reason not to violate the constraint against breaking a promise even if breaking that promise will not adversely affect anyone's utility. CSU<sub>1</sub> cannot accommodate such intuitions.

Both problems can be fixed if we abstract away from CSU<sub>1</sub>, arriving at what I call *securitist consequentialism*. In the next section, I explain securitist consequentialism and the argument for it. Since securitist consequentialism is but only a skeleton of a moral theory, I spend the following section putting some more meat on its bones, arriving at *commonsense consequentialism*. Commonsense consequentialism has all the advantages of CSU<sub>1</sub> but none of its disadvantages. It is, I believe, the most plausible version of consequentialism.

## §7.2 Securitist consequentialism and the argument for it

One of the lessons learned from the above is that we should accept *moral securitism*:

**MS** It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally

<sup>18</sup> Actually, I don't think that volitional control is necessary for being blameworthy for lacking this preference. For more on this issue, see, for instance, SCANLON 1998, A. M. SMITH 2005, and A. M. SMITH 2008.



permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ .

Moral securitism is the moral analogue of (rational) securitism (see Chapter 6), but, as with (rational) securitism, moral securitism isn't very substantive. We need, then, to decide what substantive version of moral securitism to accept. Given the arguments from the first two chapters, I think that we should accept the following teleological version of moral securitism, viz. *securitist consequentialism*:

**SC** The deontic status of a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) is determined by the reasons there are for and against the agent's preferring certain outcomes to others, such that, if S is, as of  $t_i$ , morally required to perform  $\alpha_j$ , then, of all the outcomes that S could bring about by performing some MSA that is, as of  $t_i$ , securable by S, S has most reason to desire some subset of those that would result from S's performing an MSA that includes S's performing  $\alpha_j$ .

The argument for SC runs as follows:

- 7.1 The deontic status of an act-set,  $\alpha_j$ , is determined by the agent's reasons for and against performing  $\alpha_j$ , such that, if S is, as of  $t_i$ , morally required to perform  $\alpha_j$ , then S has most reason, all things considered, to perform  $\alpha_j$ . (The Revised Version of Moral Rationalism or MR\*)
- 7.2 The agent's reasons for and against performing  $\alpha_j$  are determined by her reasons for and against preferring certain outcomes to others, such that, if S has most reason to perform  $\alpha_j$ , then, of all the outcomes that S could bring about by performing some MSA that is, as of  $t_i$ , securable by S, S has most reason to desire some subset of those that would result from S's performing an MSA that includes S's performing  $\alpha_j$ . (The Revised Version of the Teleological Conception of Practical Reasons or TCR\*)
- 7.3 Therefore, the deontic status of a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) is determined by the reasons there are for and against S's preferring certain outcomes to others, such that, if S is, as of  $t_i$ , morally required to perform  $\alpha_j$ , then, of all the outcomes that S could bring about by performing some

MSA that is, as of  $t_i$ , securable by S, S has most reason to desire some subset of those that would result from S's performing an MSA that includes S's performing  $\alpha_j$ . (SC)

This argument is similar to the one given for act-consequentialism in §1.2 and reiterated in §2.3. The original premises have, however, been revised, which accounts for the conclusion's being securitist consequentialism as opposed to act-consequentialism. Let me now explain why we should accept these revised premises, starting with 7.1 (MR\*), which is a revised version of moral rationalism (MR)—the view that an act's deontic status is determined by the reasons there are for and against performing it, such that, if S is morally required to perform  $x$ , then S has most reason (indeed, decisive reason) to perform  $x$ . The problem with MR is that it is unnecessarily narrow in that it focuses exclusively on individual actions. We should broaden it to cover sets of actions. We can do so by replacing it with MR\*. And we don't lose anything in the process, for the same considerations that support the contention that an agent can be morally required to perform an individual act,  $x$ , only if she has decisive reason to perform  $x$  also support the broader contention that an agent can be morally required to perform a set of actions,  $\alpha_j$ , only if she has decisive reason to perform  $\alpha_j$ .

To illustrate, suppose that we're wondering whether S is morally required to perform  $\alpha_j$ . Given the arguments in Chapter 2, we should hold that S is morally required to perform  $\alpha_j$  only if S has decisive reason to perform  $\alpha_j$ . To think otherwise would be to allow that S could be morally required to perform  $\alpha_j$  even though S has sufficient reason to perform some alternative—say,  $\alpha_i$ . And we should reject this possibility, for it would entail that S could be morally blameworthy for having freely and knowledgeably performed  $\alpha_i$  even though she was led to do so by flawlessly exercising her capacity to respond appropriately to the relevant reasons—the very capacity in virtue of which she is morally accountable for having performed  $\alpha_i$ . But to hold someone accountable for having performed  $\alpha_i$  in virtue of her possessing a certain capacity that, when exercised flawlessly, leads her to perform  $\alpha_i$  is inappropriate. We should, then, endorse the broader 7.1 and not just the narrower MR, for the same arguments that were given in support of MR in Chapter 2 can, *mutatis mutandis*, be given in support of 7.1 (MR\*) as well.

Let's turn now to 7.2 (TCR\*). Premise 7.2 is a revised version of the teleological conception of practical reasons (TCR)—the view according to which S has more reason to perform  $a_i$  than to perform  $a_j$  just when, and

because, S has more reason to desire that  $o_i$  obtains than to desire that  $o_j$  obtains. There are two problems with TCR. First, it is too narrow in that it focuses exclusively on individual acts; we should broaden it to encompass act-sets. Second, it's formulated on the assumption that actualism is true.<sup>19</sup> But, as I've argued above, securitism is superior to both actualism and possibilism. We should, therefore, reformulate TCR so that it's compatible with securitism, as is the case with 7.2. For these two reasons, we should accept 7.2 (TCR\*) as opposed to TCR.<sup>20</sup>

Given that we should both accept 7.1 (MR\*) as opposed to MR and accept 7.2 (TCR\*) as opposed to TCR, it follows that we should accept securitist consequentialism (SC) as opposed to act-consequentialism (AC), for 7.3 is entailed by the conjunction of 7.1 and 7.2.

### **§7.3 Commonsense consequentialism and how it compares with traditional act-consequentialism**

Although securitist consequentialism is a substantive version of moral securitism, it still isn't all that substantive as versions of consequentialism go, for it doesn't tell us how to rank outcomes. Drawing on the insights garnered from Chapter 5 and the above, we should, I believe, rank outcomes as DRAC does, although we should, unlike DRAC, distinguish between requiring and non-requiring reasons for desiring an outcome. For, as we learned above, not all moral reasons for desiring an outcome have moral requiring strength—that is, not all moral reasons for desiring an outcome are such as to make one blameworthy for lacking that desire even in the absence of countervailing reasons for not having that desire. By adopting DRAC's dual-ranking structure, we will ensure that the resulting theory can, as does DRAC, accommodate the basic features of commonsense morality, such as, supererogation, agent-centered options, and the self-other asymmetry. So, the result of conjoining securitist consequentialism with DRAC's dual-ranking method of ranking outcomes, while allowing that there are both requiring and non-requiring for desiring an outcome, is what I call *commonsense consequentialism*:

**CSC** (a) It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq$

<sup>19</sup> This is because  $o_i$  is just the outcome that *would* obtain were S to perform  $a_i$ , and likewise for  $o_j$  and  $a_j$ .

<sup>20</sup> I'm assuming that the arguments that were given in support TCR in Chapter 3 could, *mutatis mutandis*, be given in support of 7.2 as well.

$t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as to  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by S whose outcome S has both more requiring reason and more reason, all things considered, to want to obtain.

Clause *a* ensures that CSC is a version of moral securitism. Clause *b* ensures that it is a maximizing, teleological version of moral securitism. And clause *c* ensures that it takes the ranking of outcomes to be function of two auxiliary rankings, just as DRAC does.

Since CSC is a bit unwieldy, it might be useful to think of CSC as a three-step procedure (although it is, of course, a criterion of rightness, not a decision procedure):

- Step 1: Survey all the various MSAs that are, as of  $t_i$ , securable by S.
- Step 2: For each of these MSAs, determine whether or not that MSA is objectively morally permissible. A given MSA is objectively morally permissible if and only if the agent, S, has both optimal requiring reason and optimal reason, all things considered, to want its outcome to obtain.<sup>21</sup>
- Step 3: Judge that S is, as of  $t_i$ , objectively morally permitted to perform  $\alpha_j$  if and only if one of the objectively morally permissible MSAs that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ .

To acquire a better understanding of CSC, it will be helpful to compare it with *traditional act-consequentialism* and its *criterion of rightness*:

<sup>21</sup> Recall that S has optimal reason to  $\varphi$  if and only if there is no alternative,  $\psi$ , such that S has more reason to  $\psi$  than to  $\varphi$ .

**TAC<sub>cr</sub>** An act,  $x$ , is morally permissible if and only if it maximizes the good (impersonally construed).<sup>22</sup>

As I explain below, CSC departs from TAC in five important respects. The first two concern how the two theories differ in their criteria for determining moral permissibility, the next two concern how they differ in their methods for ranking outcomes, and the last concerns how they differ in their abilities to accommodate certain basic features of commonsense morality.

§§7.3.1 *Whereas TAC offers a criterion for evaluating only individual actions, CSC offers a criterion for evaluating sets of actions:* TAC<sub>cr</sub> is a criterion for evaluating only individual actions (types and tokens). Unfortunately, though, not all sets of actions constitute larger individual actions. And, thus, TAC<sub>cr</sub> is unable to evaluate certain sets of actions. Yet, as we've seen, this is something that any adequate moral theory must be able to do if it's going to account for many of our typical moral judgments, as we often do make moral judgments concerning various sets of actions. It's fortunate, then, that CSC avoids this problem, for, unlike TAC<sub>cr</sub>, CSC's criterion of rightness is formulated using the variable,  $\alpha_j$ , which ranges over sets of actions.

It is because CSC can assess the deontic statuses of sets of actions that it is compatible with the idea that there are imperfect duties, duties that can be fulfilled only by performing certain sets of actions over time. For instance, you can't fulfill the duty of beneficence by performing any single beneficent act. To fulfill this duty, you must act in such a way as to demonstrate that you have adopted the happiness of others as a major, *on-going*, and life-shaping end, and the only way to do that is to complete some fairly extended set of actions that demonstrates an on-going commitment to promoting the welfare of others.

§§7.3.2 *Whereas TAC is a form of direct consequentialism, CSC is a form of indirect consequentialism:* CSC is a version of securitist consequentialism,

<sup>22</sup> TAC<sub>cr</sub> is not equivalent to traditional act-consequentialism, for one can both accept TAC<sub>cr</sub> and reject act-consequentialism, but one cannot both accept traditional act-consequentialism and reject act-consequentialism. The reason that one cannot both accept traditional act-consequentialism and reject act-consequentialism is that traditional act-consequentialism is, by definition, a species of act-consequentialism; traditional act-consequentialism holds both that act-consequentialism is true and that an act produces the outcome that the agent has most reason to desire if and only if it maximizes the good. One can, by contrast, accept TAC<sub>cr</sub> and reject act-consequentialism, as where one denies that the outcome that a given agent has most reason to desire is always the one in which the good is maximized.

which is a form of *indirect consequentialism*. Whereas, on *direct consequentialism* (i.e., act-consequentialism), the deontic status of an action is determined by how its outcome ranks relative to those of the available alternatives on some evaluative ranking of outcomes, on *indirect consequentialism*, the deontic status of an action is determined by whether or not it accords with some ideal set of rules (rule-consequentialism) or stems from some ideal set of motives (motive-consequentialism) or is contained within some ideal set of MSAs (securitist consequentialism). The ideal set of rules, motives, or MSAs is in turn selected on the basis of how their associated outcomes rank relative to those of the available alternatives on some evaluative ranking of outcomes. Hence, the consequences are only indirectly relevant in assessing the deontic statuses of actions.

To illustrate, consider rule-consequentialism. It holds that an act is morally permissible if and only if it accords with the ideal set of rules. And, on at least one version of rule-consequentialism, the ideal set of rules is the one “whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being (with some priority for the worst off)” (HOOKER 2000, p. 32). Similarly, securitist consequentialism holds that an act is morally permissible if and only if it is contained within the ideal set of MSAs. And, on securitist consequentialism, the ideal set of MSAs (i.e., the permissible MSAs) are those whose outcomes are not, on the agent’s evaluative ranking, outranked by that of any alternative MSA securable by the agent.

Being a version of securitist consequentialism as opposed act-consequentialism is what enables CSC to endorse moral securitism: the view that it is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S’s performing  $\alpha_j$ . As I argued above, we should endorse moral securitism given the problems associated with its two main rivals: actualism and possibilism.

§§7.3.3 *Whereas TAC ranks outcomes along a single evaluative dimension, CSC ranks outcomes according to a function of two auxiliary rankings, each along a different evaluative dimension:* If moral rationalism (MR) is true, then it is always morally permissible to do what there is most reason to do, all things considered. This and the fact that moral reasons are not rationally overriding implies that they’re not morally overriding either—that is, that non-moral reasons can, and sometimes do, prevent moral reasons, even those with substantial moral requiring strength, from generating moral requirements. Thus, non-moral reasons serve to limit what we can be

morally required to do. If, for instance, the non-moral reasons that you have to perform some self-interested act,  $x$ , outweigh the moral reasons that you have to perform some altruistic alternative act,  $y$ , then you can't be morally required to sacrifice your self-interest and perform  $y$  as opposed to  $x$ . (Of course, you wouldn't be morally required to perform  $x$  either, for surely you would be permitted to do what you have more moral reason to do rather than what you have less moral reason to do.)

The upshot of all this is that the deontic status of an action is a function of both moral reasons and non-moral reasons. That function, I've argued in Chapter 5, is:

**MP**  $S$ 's performing  $x$  is *morally permissible* if and only if there is no available alternative that  $S$  has both more requiring reason and more reason, all things considered, to perform (where a requiring reason is just a moral reason with some moral requiring strength).

And, in that same chapter, I argued that the only way to accommodate MP within a consequentialist framework is to hold that one outcome outranks another, on the agent's evaluative ranking, if and only if she has both more moral (requiring) reason and more reason, all things considered, to want the one to obtain than to want the other to obtain. Thus, on CSC, the ranking of outcomes is a function of two auxiliary rankings: one a ranking in terms of how much requiring reason the agent has to desire that they obtain and the other in terms of how much reason, all things considered, the agent has to desire that they obtain.

As I showed in Chapter 4, this dual-ranking structure allows the consequentialist to accommodate agent-centered options (of both the agent-favoring and the agent-sacrificing varieties), for it allows the consequentialist to give agents the option to either discount their own interests entirely (as in the case of agent-sacrificing options) or to give their interests much more weight than the equivalent interests of others (as in the case of agent-favoring options).  $TAC_{cr}$ , by contrast, ranks outcomes along a single evaluative dimension: their impersonal value. Thus, unlike CSC, there is no way for  $TAC_{cr}$  to accommodate options to give one's own interests anything but the weight that they carry from the impersonal point of view.

§§7.3.4 *Whereas TAC ranks outcomes according to their value, CSC ranks outcomes according to how much reason the agent has to desire that they obtain:* Suppose that I can effect only one of two possible worlds: either the one in which I save my own daughter (call this  $w_1$ ) or the one in which I ensure

that some stranger saves his slightly more gifted daughter (call this  $w_2$ ). And let me just stipulate that  $w_2$  is slightly better than  $w_1$ , as the stranger's daughter would contribute slightly more impersonal good to the world than my daughter would. We'll assume, though, that everything else is equal. Now, despite the fact that  $w_2$  is slightly better than  $w_1$ , it seems that I have more reason to desire that  $w_1$  be actualized than to desire that  $w_2$  be actualized, for it seems that the reason that I have to desire the world in which my daughter is saved outweighs the reason that I have to desire the world in which there is slightly more impersonal good. And it's not just that I have more reason, *all things considered*, to desire that  $w_1$  be actualized. Given the moral importance of the special relationship that I have with my daughter, I also have more *requiring* reason to desire that  $w_1$  be actualized. Indeed, if I didn't care more about the preservation of my daughter's life than I did about such small increases in the impersonal good, I would be open to moral criticism. Fathers ought, morally speaking, to care more about the preservation of their children's lives than they do about such small increases in the impersonal good.

If all this is right, then CSC and TAC will sometimes rank these two worlds differently. On TAC, there is only one possible ranking of these two worlds, and, on that ranking,  $w_2$  outranks  $w_1$ . On CSC, by contrast, there is potentially a different ranking for each agent, as each agent is to rank outcomes according to how much reason she has to desire that these outcomes obtain. Since one agent can have more reason than another to desire that a certain outcome obtains, different agents can end up with different rankings. This is true in the above example. On my evaluative ranking,  $w_1$  outranks  $w_2$ , but, on the stranger's evaluative ranking,  $w_2$  outranks  $w_1$ . Thus, whereas TAC holds that we all have the same goal of bringing about the best possible outcome, CSC holds that each individual agent has the potentially distinct goal of bringing about the outcome that *she* has most reason to desire. It is because CSC, unlike TAC, holds that each agent has her own ranking of outcomes that CSC, unlike TAC<sub>cr</sub>, can accommodate agent-centered constraints, such as the special obligation that parents have to ensure that they save their own children even at the cost of fewer children being saved by their parents.

The problem with ranking outcomes according to their value, then, is that it excludes from consideration all agent-relative reasons. This is because the value of a possible world correlates with only a subset of an agent's reasons for desiring its actualization—specifically, a subset of her



agent-neutral reasons for desiring its actualization.<sup>23</sup> But agents can have agent-relative reasons for desiring its actualization, as is the case with my reasons for preferring  $w_1$ —the possible world in which my daughter is saved. Moreover, these agent-relative reasons can tip the balance such that agents sometimes have better overall reason to prefer a worse world to a better world, as is the case where I have better overall reason to prefer  $w_1$  to  $w_2$ . And, as I argued in Chapter 3, if I have better overall reason to prefer  $w_1$  to  $w_2$ , then I have better overall reason to *do* what will actualize  $w_1$  as opposed to what will actualize  $w_2$ . Given this and the arguments for moral rationalism from Chapter 2, we should conclude, contrary to what TAC<sub>cr</sub> implies, that I cannot be obligated to act so as to actualize  $w_2$  as opposed to  $w_1$  even though  $w_2$  has more impersonal value. The lesson, then, is that if we want to take into account all of our reasons for desiring a given outcome and, thereby, all our reasons for acting so as to produce that outcome, we must rank outcomes according to how much reason the agent has to desire that they obtain as opposed to how valuable they are.

§§7.3.5 *Whereas TAC cannot accommodate the basic features of commonsense morality, CSC can:* CSC can accommodate all the basic features of commonsense morality: agent-centered restrictions, special obligations, agent-favoring options, agent-sacrificing options, and supererogatory acts. Let me briefly illustrate each in turn. To account for an agent-centered restriction against, say, the commission of murder, the proponent of CSC need only claim that, other things being equal, S has both more requiring reason and more reason, all things considered, to prefer the outcome of an MSA in which she refrains from murdering some individual to the outcome of an otherwise identical MSA in which she instead murders that individual to prevent less than  $n$  others from each committing a comparable murder—where  $n$  marks the threshold for this non-absolute restriction.<sup>24</sup> To account for the special obligation that, say, parents have to

<sup>23</sup> It's only a subset of her reasons, because we must not only meet the partiality challenge, but also the wrong-kind-of-reason problem. See Chapter 3.

<sup>24</sup> Note the ease with which CSC can account for the fact that agent-centered restrictions have thresholds: at some point, the agent-neutral reason that one has to want there to be fewer infringements of the restriction overall outweigh the agent-relative reason that one has to want there to be no personal infringements of that restriction. Deontological views have a much more difficult time accounting for thresholds—see, for instance, ELLIS 1992. CSC can with similar ease account for why some agent-centered restrictions are more stringent than others, such that one can permissibly infringe upon one but not the other in instances where the two conflict. The proponent of CSC can claim, for instance, that the agent-relative reason that one has to refrain from committing murder outweighs, other things being equal, the agent-relative reason one has to refrain from breaking a promise.

nurture their own children, the proponent of CSC need only claim that, other things being equal, S has both more requiring reason and more reason, all things considered, to prefer the outcome of an MSA in which she nurtures her own children to the outcome of an otherwise identical MSA in which she instead abandons her own children so as to ensure that  $n$  others nurture their children—where  $n$  marks the threshold for this non-absolute special obligation.

To account for the agent-favoring option that S currently has, say, to perform either an altruistic act or a self-interested act, the proponent of CSC need only claim that there is both a morally permissible MSA available to S in which she presently performs an altruistic act and a morally permissible MSA available to her in which she presently performs a self-interested act, for, in either case, a reasonable balance between altruism and egoism will be achieved. To account for the agent-sacrificing option that S currently has, say, to use her last dose of painkillers so as to alleviate either some stranger's mild headache or her own more severe headache, the proponent of CSC need only claim that, other things being equal, S has more requiring reason to prefer the outcome of an MSA in which she does the former, but more reason, all things considered, to prefer the outcome of an otherwise identical MSA in which she instead does the latter. And, to account for the fact that it would be supererogatory for her to do the former, the proponent of CSC need only both accept that S has more moral reason to prefer the outcome of an MSA in which she uses her last dose of painkillers to alleviate the stranger's mild headache and accept the account of supererogation that I gave above. We see, then, that CSC shares with DRAC the ability to accommodate these basic features of commonsense morality.

#### **§7.4 What has been shown and what remains to be shown**

I've argued for a rationalist conception of the authority of morality, for a teleological conception of practical reasons, for a teleological conception of morality, and for a teleological conception of (objective) rationality. Respectively, these views are:

- MR\***            The deontic status of an act-set,  $\alpha_j$ , is determined by the agent's reasons for and against performing  $\alpha_j$ , such that, if S is, as of  $t_i$ , morally required to perform  $\alpha_j$ , then S has most reason, all things considered, to perform  $\alpha_j$ .

- TCR\*** The agent's reasons for and against performing  $\alpha_j$  are determined by her reasons for and against preferring certain outcomes to others, such that, if S has most reason to perform  $\alpha_j$ , then, of all the outcomes that S could bring about by performing some MSA that is, as of  $t_i$ , securable by S, S has most reason to desire some subset of those that would result from S's performing an MSA that includes S's performing  $\alpha_j$ .
- CSC** (a) It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as to  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by S whose outcome S has both more requiring reason and more reason, all things considered, to want to obtain.
- TMS** (a) It is, as to  $t_i$ , objectively rationally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively rationally permissible maximal sets of actions that are, as to  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively rationally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as to  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as to  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as to  $t_i$ , securable by S whose outcome S has more reason, all things considered, to want to obtain.

I've shown that these four theses form a coherent whole that together provide a systematic account of our commonsense intuitions about morality, rationality, and the relationship between the two. MR\*, for instance, accounts for our commonsense intuition that morality, unlike the law, is rationally authoritative. TCR\* systemically accounts for our commonsense intuitions about what we have reason to do. CSC systematically accounts for our commonsense intuitions about what we (objectively) morally ought to do, providing, in the process, a systematic account of all the basic features of commonsense morality: supererogation, imperfect duties, special obligations, agent-favoring options, agent-sacrificing options, agent-centered restrictions, and the self-other asymmetry. Lastly, TMS systematically accounts for our commonsense intuitions about what we (objectively) rationally ought to do, providing, in the process, what is perhaps the most plausible explanation for the basic belief—the belief that, in most typical choice situations, the relevant reasons do not require performing some particular alternative, but instead permit performing any of numerous alternatives.

I've shown that maximizing, teleological theories—whether they be about reasons, morality, or rationality—are extremely ecumenical and accommodating. These theories can accommodate our commonsense intuitions about reasons, morality, and rationality while remaining compatible with a wide range of theses about the nature of value, morality, and persons—even those that have seemed to some (e.g., Anderson and Scanlon) antithetical to the teleological approach. I've shown, for instance, that these theses are compatible with the view that the primary bearers of value are not states of affairs but instead persons, animals, and things, and that our basic evaluative attitudes—attitudes such as love, respect, and admiration—are non-propositional attitudes that we take up toward persons, animals, and things. Surprisingly, then, we find that these teleological theories are compatible with even the idea that persons, not states of affairs, have intrinsic value and that we respond appropriately to their value by loving, respecting, and admiring them. Moreover, these theories are compatible with the view that we are morally prohibited from acting in ways that fail to express our respect for persons even when our doing so would prevent more numerous others from acting in ways that are comparably disrespectful to persons. What's more, these theories have the ability to surpass some of its non-teleological rivals in accommodating

our commonsense intuitions, doing a better job of accommodating and accounting for our commonsense intuitions concerning intra-agent cases.<sup>25</sup>

Admittedly, though, these teleological views (i.e., TCR\*, CSC, and TMS) will be compatible with our commonsense intuitions only if certain views about what agents have reason to desire are correct. The real question, then, is whether what agents have reason to desire is such that, when inputted into these teleological theories, the output is plausible deontic verdicts. I've suggested that this is so. I've suggested, for instance, that it's intuitive to think that, other things being equal, an agent ought to prefer the state of affairs in which she saves her loved one to the state of affairs in which she enables some stranger to save her loved one, even if the latter state of affairs would be slightly better, impersonally speaking. We would, I think, fault an agent for lacking such a preference, thinking either that she is not responding appropriately to the reasons that their loving relationship provides or that she must not truly love her putative "loved" one. And, as I argued in Chapter 4, this view about what agents have reason to desire is precisely the one that the consequentialist needs to ground the special obligation that agents have to save their own loved ones.

I also argued, again in Chapter 4, that even in those cases in which our pre-theoretical judgments about what we have reason to desire are not quite what's needed to generate plausible deontic verdicts when inputted into a teleological theory such as CSC, we might, on reflection, come to a different judgment about what there is reason to desire in light of the intuitive attractiveness of MR\* and TCR\* and their entailment of securitist consequentialism (where CSC is, perhaps, the most plausible version of securitist consequentialism). Of course, all this is highly speculative. Ideally, what we want is a fully defended and developed theory about what agents have reason to desire and to be shown that such a theory, when conjoined with theories such as CSC and TMS, yields plausible deontic verdicts. Unfortunately, I have no such theory to offer. There will, then, be some skepticism as to whether these teleological theories will in fact do all that I say that they can. That's understandable.

I should remind my readers, though, that my goal has not been to convince them that CSC is the most plausible moral theory or that TMS is the most plausible theory of objective rationality. To have any hope of succeeding in such ambitious goals, I would have to provide a fully defended and developed theory about what agents have reason to desire.

<sup>25</sup> See §4.3.

What's more, I would need to survey all their rivals and do a thorough assessment of how CSC and TMS compare to them. I've done neither. But, then, my goal has been considerably less ambitious. I hope only to have developed the most plausible versions of these theories and to have shown that these theories are attractive and promising and so deserve our further consideration. The advantages and disadvantages of theories that take the deontic statuses of actions to be a function of the value of their outcomes have been well explored, but the same cannot be said of the advantages and disadvantages of theories that take the deontic statuses of actions to be a function of how much reason agents have to want their outcomes to obtain—theories such as DRAC, CSC, and TMS. My hope is to have got this exploration underway and to have uncovered enough to demonstrate that such theories merit further exploration.

## Glossary

**Alternative:** Alternatives are *agent-identical*, *time-identical*, mutually exclusive, and jointly exhaustive.

**Across-the-board approach:** On this approach, we assess permissibility by directly applying normative principles to *minimal acts*, *maximal sets of actions*, and everything in between. Cf. *bottom-up approach* and *top-down approach*.

**Act-consequentialism (AC)—a.k.a. rational-desire teleology:** This is the view that an act's *deontic status* is determined by the agent's reasons for and against preferring its *outcome* to those of the *available* alternatives, such that, if S is morally required to perform *x*, then, of all the outcomes that S could bring about, S has most (indeed, decisive) reason to desire that *x*'s outcome obtains.

**Action:** An action is something that an agent does because she intends to do it. Actions include *act-types*, *act-tokens*, and *act-sets*.

**Act-set:** An act-set is a set of one or more acts (types or tokens) that are all jointly performable by a single agent. The acts in the set may be simple or compound, synchronous or asynchronous, consecutive or inconsecutive. I'll use  $\alpha_1, \alpha_2, \dots, \alpha_n$  to designate act-sets.

**Act-token:** An act-token is a particular act performed by a particular agent at a particular time. Act-tokens are performable at most once and are not the sort of thing that can be instantiated. I'll use  $a_1, a_2, \dots, a_n$  to designate act-tokens. And I'll use  $x$  (or  $x_i$ ) as a variable that ranges over individual acts, whether they be types or tokens.

**Act-type:** An act-type is universal that can, in general, be instantiated by a number of distinct *act-tokens*. I'll use  $A_1, A_2, \dots, A_n$  to designate act-types.

And I'll use  $x$  (or  $x_i$ ) as a variable that ranges over individual acts, whether they be types or tokens.

**Actualism:** This is the view according to which it is, as to  $t_i$ , permissible for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if the *set of actions* that S would actually perform were S to perform  $\alpha_j$  is no worse than the set of actions that S would actually perform were S to perform any *alternative* set of actions that is, as to  $t_i$ , *personally possible* for S. Cf. *possibilism* and *securitism*.

**Agent-centered constraints:** An agent-centered constraint is a constraint on maximizing the good that it would be wrong to *infringe* upon even in some circumstances where doing so would minimize comparable infringements of that constraint. (Note, though, that constraints don't have to be absolute. It may be permissible to infringe upon a constraint if enough good is at stake, and it may be permissible to infringe upon a constraint if doing so would prevent a sufficient number of comparable infringements of that constraint.) There are two types of agent-centered constraints: *agent-centered restrictions* and *special obligations*.

**Agent-centered options:** An agent-centered option is a *moral option* either to act so as to make things better overall but worse for oneself (or others) or to act so as to make things better for oneself (or others) but worse overall. These options provide agents with the freedom to give their own interests more or less weight than they have from the impersonal perspective. There are two types of agent-centered options: *agent-favoring options* and *agent-sacrificing options*.

**Agent-centered restrictions:** An agent-centered restriction is a type of *constraint*. There is an agent-centered restriction against performing a certain *act-type* (such as, murder) if and only if agents are prohibited from performing that act-type even in some circumstances in which performing that act-type is the only way to minimize comparable commissions of that act-type.

**Agent-favoring options:** An agent-favoring option is a type of *agent-centered option*. Agent-favoring options are *moral options* either to promote one's own interests or to sacrifice one's own interests for the sake of doing more to promote the overall good.



**Agent-identical:** Two sets of actions are agent-identical if and only if they have the same agent.

**Agent-sacrificing options:** An agent-sacrificing option is a type of *agent-centered option*. Agent-sacrificing options are *moral options* either to maximize the impersonal good or to sacrifice one's own self-interests for the sake of doing more to promote the interests of others even when doing so is not what's best overall, as where the extent of one's own self-sacrifice would exceed the extent of the benefits thereby bestowed upon others.

**Agglomeration Principle:** According to the following agglomeration principle, permissibility agglomerates over conjunction:  $[P(S, t_i, x_1), P(S, t_i, x_2), \dots, \& P(S, t_i, x_n)] \rightarrow P(S, t_i, [x_1, x_2, \dots, \& x_n])$ , where 'P(S,  $t_i$ ,  $\varphi$ )' stands for 'S is, as of  $t_i$ , permitted to  $\varphi$ '.

**Aggregate utility:** The aggregate *utility* produced by an act is the sum of all the utility it produces minus the sum of all the disutility it produces.

**Asymmetric relation:** A relation, R, is asymmetric just in case: if  $xRy$ , then  $\sim yRx$ .

**Available:** Roughly speaking,  $\varphi$  is available to S if and only if S can  $\varphi$ —that is, if and only if  $\varphi$ -ing is *personally possible* for S.

**The basic belief:** This is the belief that, in most typical choice situations, the relevant reasons do not require performing some particular *alternative*, but instead permit performing any of numerous alternatives.

**Best alternative:** The best *alternative* is the one that the agent has most reason to perform. Cf. *morally best*.

**Better for:** For all subjects S and all states of affairs  $p$  and  $q$ , it is better for S that  $p$  is the case than that  $q$  is the case if and only if the set of all the right kind of reasons to prefer its being the case that  $p$  to its being the case that  $q$  for S's sake is weightier than the set of all the right kind of reasons to prefer its being the case that  $q$  to its being the case that  $p$  for S's sake.

**Better K:** For all kinds K and all things A and B of kind K, A is a better K than B is if and only if the set of all the right kind of reasons to prefer A to B

when selecting a  $K$  is weightier than the set of all the right kind of reasons to prefer  $B$  to  $A$  when selecting a  $K$ .

**Better that:** For all states of affairs  $p$  and  $q$ , it is better that  $p$  is the case than that  $q$  is the case if and only if the set of all the right kind of reasons to prefer its being the case that  $p$  to its being the case that  $q$  is weightier than the set of all the right kind of reasons to prefer its being the case that  $q$  to its being the case that  $p$ .

**Bottom-up approach:** On this approach, we assess permissibility by directly applying normative principles to only *minimal acts* and then *agglomerating* to determine the permissibility of *sets of action*. Cf. *across-the-board approach* and *top-down approach*.

**Blameworthy:**  $S$  is blameworthy for performing  $x$  if and only if it is appropriate (i.e., fitting) for  $S$  to feel guilt about having performed  $x$  and appropriate for others to feel indignation in response to  $S$ 's having performed  $x$  and, if they were thereby wronged, to resent  $S$  for having performed  $x$ .

**Commonsense consequentialism (CSC):** (a) It is, as of  $t_i$ , objectively morally permissible for  $S$  to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by  $S$  includes  $S$ 's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on  $S$ 's evaluative ranking, outranked by that of any alternative MSA that is, as to  $t_i$ , securable by  $S$ . (c)  $MSA_i$ 's outcome is not, on  $S$ 's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by  $S$  if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by  $S$  whose outcome  $S$  has both *more requiring reason* and *more reason*, all things considered, to want to obtain.

**Commonsense utilitarianism (CSU):** (a) It is, as of  $t_i$ , objectively morally permissible for  $S$  to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by  $S$  includes  $S$ 's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on  $S$ 's evaluative ranking, outranked by that of any alternative MSA

that is, as of  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by S whose outcome contains both more constraint-adjusted utility and more comprehensively-adjusted utility than  $MSA_i$ 's outcome does, where the constraint-adjusted utility of an outcome is just the sum of the utility for others, adjusted by multiplying any disutility resulting from S's infringements of an agent-centered constraint by 500, and where the comprehensively-adjusted utility of an outcome is just its constraint-adjusted utility added to the product of S's utility times ten.

**Compound act:** A compound act is an act that has two or more simpler acts as parts—e.g., assembling a model airplane and chewing gum while walking. Cf. *simple act*.

**Consequentialism:** See *act-consequentialism*.

**Consequentialist prudence:** This is the view that S's performing  $x$  is prudent if and only if it *maximizes* S's utility.

**Consequentialize:** To consequentialize a nonconsequentialist theory is to construct a substantive version of *consequentialism* that is *deontically equivalent* to that nonconsequentialist theory.

**Constraining right:** By definition, a potential victim, V, has a constraining right against being  $\varphi$ -ed (e.g., murdered) if and only if V has a right not to be  $\varphi$ -ed even in some circumstances where the agent's  $\varphi$ -ing V would minimize comparable commissions of  $\varphi$  by herself or others.

**Constraint:** See *agent-centered constraint*.

**Counterfactual determinism:** This is the view that, for each set of actions that S might perform, there is some determinate fact as to what the world would be like were S to perform that set of actions and that, for each set of intentions that S might form, there is some determinate fact as to what the world would be like were S to form those intentions.

**Decisive reason:** S has decisive reason to  $\varphi$  if and only if S's reasons are such as to make S *objectively rationally* required to  $\varphi$ . In other words, S has decisive reason to  $\varphi$  if and only if S does not have *sufficient reason* to  $\sim\varphi$ .

**Defeat:** To say that the reasons that favor  $\varphi$ -ing defeat the reasons that favor  $\sim\varphi$ -ing is to say that the reasons that favor  $\varphi$ -ing prevail over the reasons that favor  $\sim\varphi$ -ing such that the agent has *decisive reason* to  $\varphi$ .

**Deontically equivalent:** Two theories are deontically equivalent if and only if they agree about the deontic status of every (actual or possible) act—that is, for any deontic predicate (such as, *permissible*, *impermissible*, *optional*, *obligatory*, or *supererogatory*), both theories are in perfect agreement as to the set of actions that are in the extension of that predicate.

**Deontic equivalence thesis (DET):** This is the thesis that, for any remotely plausible *nonconsequentialist* theory, there is a version of *consequentialism* that is *deontically equivalent* to it.

**Deontic status:** All acts have one of the following two deontic statuses: (1) *permissible* or (2) *impermissible*. All permissible acts additionally have one of the following two deontic statuses: (1.1) *optional* or (1.2) *obligatory*. And all optional acts additionally have one of the following two deontic statuses: (1.1.1) *merely permissible* or (1.1.2) *supererogatory*.

**Deontic moral value:** The deontic moral value of an act is a measure of how much (objective) moral reason there is to perform it.

**Deontic value:** The deontic value of an act is a measure of how much (objective) reason there is to perform it.

**Deontological prudence:** This is the view that there are certain acts that are intrinsically imprudent such that it would be imprudent to perform them even if doing so is clearly what would best promote the agent's *utility* over time.

**Deontology:** A moral theory is deontological if and only if it is *nonconsequentialist* and includes at least one *agent-centered constraint*.

**Direct consequentialism (a.k.a. *act-consequentialism*):** On direct consequentialism, the *deontic status* of an action is determined by how its *outcome* ranks relative to those of the available alternatives on some *evaluative ranking* of outcomes. Cf. *indirect consequentialism*.

**Distribution principle:** According to the following distribution principle, permissibility distributes over conjunction:  $P(S, t_i, [x_1, x_2, \dots, \& x_n]) \rightarrow [P(S, t_i, x_1), P(S, t_i, x_2), \dots, \& P(S, t_i, x_n)]$ , where 'P(S, t<sub>i</sub>, φ)' stands for 'S is, as of t<sub>i</sub>, permitted to φ'.

**Dual-ranking act-consequentialism (DRAC):** S's performing *x* is morally permissible if and only if, and because, there is no available act alternative that would produce an outcome that S has both more *moral reason* and more reason, all things considered, to want to obtain than to want *x*'s outcome to obtain.

**Egoism:** See *ethical egoism*.

**Ethical egoism:** This is the view that holds both that *act-consequentialism* is true and that an act produces the outcome that S has the most reason to desire if and only if it *maximizes* S's utility.

**Evaluative ranking:** An evaluative ranking of *outcomes* is a ranking in terms of the agent's reasons (or some subset of her reasons—e.g., her *fittingness reasons*) for preferring each outcome to the available alternatives.

**Expectably best:** An act is expectably best if and only if it has the highest expected *deontic value*. The deontic value of an act is a measure of how much reason there is to perform it. And we get the expected deontic value of an action "by multiplying the subjective probability that some practical comparative is true by the [deontic] value of that action if it is true, doing the same for all other practical comparatives, and adding up the results" — practical comparatives are anything of the form: the balance of reasons favors S's doing *x* as opposed to *y* (SEPIELLI 2009, pp. 7 & 11). Cf. *expectably morally best*.

**Expectably morally best:** An act is expectably morally best if and only if it has the highest expected *deontic moral value*. The deontic moral value of an act is a measure of how much moral reason there is to perform it. And we get the expected deontic moral value of an action by multiplying the subjective probability that some practical moral comparative is true by the deontic moral value of that action if it is true, doing the same for all other practical moral comparatives, and adding up the results—practical moral comparatives are anything of the form: the balance of moral reasons favors

S's doing  $x$  as opposed to  $y$  (Cf. SEPIELLI 2009, pp. 7 & 11). Cf. *expectably best*.

**Explanatory reasons for action:** These are the facts that explain why the agent performed the act in question. Cf. *normative reasons for action* and *motivating reason for action*.

**Fittingness reasons:** These are all and only those reasons that are relevant to determining whether, and to what extent, an outcome is valuable/desirable (i.e., fitting to desire).

**Foot's thesis (FT):** An act's outcome is better than those of the available alternatives if and only if, and because, the act that produces it is morally *obligatory*.

**Freely performs:** S freely performs  $x$  if and only if S performs  $x$  under conditions that make it appropriate (in the sense of being apt, fitting, or correct) to respond to S's having performed  $x$  with various reactive attitudes such as blame and praise.

**Good:** A state of affairs,  $p$ , is good if and only if it is *better* than most of the states of affairs in some contextually supplied comparison class.

**Good for:** A state of affairs,  $p$ , is good for a subject, S, if and only if it is *better for* S than most of the states of affairs in some contextually supplied comparison class.

**Good of a kind:** A thing,  $A$ , of kind K is a good K if and only if it is *better K* than most of things of kind K in some contextually supplied comparison class.

**Imperfect reason:** These are reasons that do not support performing any specific alternative, but instead support performing any of the alternatives that would each constitute an equally effective means of achieving the same worthy end.

**Impermissible:** S's  $\varphi$ -ing is impermissible if and only if S's  $\varphi$ -ing is not *permissible*.

**Impersonal-value teleology:** This is the view an act's *deontic status* is determined by the impersonal value of its outcome, such that, if S is morally required to perform  $x$ , then S's performing  $x$  would maximize the good. Cf. *personal-value teleology* and *rational-desire teleology*.

**Includes:** One set of actions,  $\alpha_i$ , *includes* another,  $\alpha_j$ , if and only if every element in  $\alpha_i$  is also an element in  $\alpha_j$ —that is, if and only if  $\alpha_i \subseteq \alpha_j$ .

**Indirect consequentialism:** On indirect consequentialism, the *deontic status* of an action is determined by whether or not it (a) accords with the ideal set of rules (rule-consequentialism), (b) stems from the ideal set of motives (motive-consequentialism), or (c) is included in one of the ideal set of MSAs (*securitist consequentialism*), where the ideal set of rules, motives, or MSAs is in turn selected on the basis how their associated outcomes rank relative to those of the alternatives on some evaluative ranking. Cf. *direct consequentialism*.

**Infringement:** If there is a *constraint* against performing a certain act-type, then any commission of an act of that type constitutes an infringement of that constraint. Cf. *violation*.

**Judgment-sensitive attitude:** A judgment-sensitive attitude is an attitude that is sensitive to one's judgments about reasons for and against it.

**Knowledgeably performs:** S knowledgeably performs  $x$  if and only if S performs  $x$  knowing all the relevant facts—the relevant facts being the facts the ignorance of which would otherwise either inculpate or exculpate S for performing  $x$ .

**Maximal set of actions (MSA):** A set of actions,  $\alpha_j$ , that is, as of  $t_i$ , *securable* by S is a maximal set if and only if there is no set of actions that is, as of  $t_i$ , securable by S that includes  $\alpha_j$  as a *proper subset*.

**Maximizes:** An act maximizes *utility* (or value) if and only if there is no alternative that produces more utility (or value) than it does. Note, then, that more than one alternative can maximize utility (or value).

**Maximizing act-utilitarianism (MAU):** This is the view according to which both (1) *traditional act-consequentialism* is true and (2) an act *maximizes* the good if and only if it *maximizes utility*.

**Maximizing securitist utilitarianism (MSU):** (a) It is, as of  $t_i$ , objectively morally permissible for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively morally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as of  $t_i$ , securable by S whose outcome contains more aggregate utility than  $MSA_i$ 's outcome does.

**Merely permissible:** S's  $\varphi$ -ing is merely permissible if and only if S's  $\varphi$ -ing is *optional* but not *supererogatory*.

**Meta-criterion of rightness:** Assuming that the *deontic status* of an action is a function of both moral and non-moral reasons, then a meta-criterion of rightness is what specifies what that function is, providing both necessary and sufficient conditions for an act's being morally permissible in terms of both moral and non-moral reasons.

**Minimal act:** "A *minimal* action is one which, once begun, cannot be stopped by its agent short of completion: once initiated a minimal action is to this extent out of an agent's control. All instantaneous actions are necessarily minimal: for example, ...placing a bet. And some non-instantaneous actions are contingently minimal: for example, ...beheading by guillotine" (Sobel 1976, p. 198). Cf. *compound act*.

**Moral dilemma:** A "moral dilemma arises when an agent is in a choice situation in which he/she cannot satisfy the dictates of morality" (VALLENTYNE 1989, 301). See also *obligation dilemma* and *prohibition dilemma*.

**Moral justifying strength:** One reason,  $R_1$ , has more moral justifying strength than another,  $R_2$ , if and only if:

- (i)  $R_1$  would make it morally permissible to do anything that  $R_2$  would make it morally permissible to do.



- (ii)  $R_1$  would make it morally permissible do some things that  $R_2$  would *not* make it morally permissible to do.

**Morally best alternative:** The morally best *alternative* is the one that the agent has most *moral reason* to perform. Cf. *best*.

**Morally overriding:** Moral reasons are morally overriding if and only if S has more moral reason to perform  $x$  than to perform  $y$  ( $y \neq x$ ) only if S is not morally permitted to perform  $y$ . Cf. *rationally overriding*.

**Morally relevant reason:** Any reason that is relevant to determining an act's deontic status is a morally relevant reason. This includes any *non-moral reasons with moral justifying strength*. Cf. *moral reason*.

**Moral option:** A moral option exists whenever an agent is morally permitted to perform more than one *alternative*.

**Moral rationalism (MR):** An act's *deontic status* is determined by the agent's reasons for and against performing it, such that, if a subject, S, is morally required to perform an act,  $x$ , then S has most (indeed, decisive) reason to perform  $x$ . See also the *revised version of moral rationalism*.

**Moral reason:** A moral reason to  $\varphi$  is a reason that, morally speaking, counts in favor of  $\varphi$ -ing. Such a reason would, if unopposed and sufficiently weighty, be capable of making  $\varphi$  either morally obligatory or morally supererogatory. Cf. *morally relevant reason*.

**Moral requiring strength:** One reason,  $R_1$ , has more moral requiring strength than another,  $R_2$ , if and only if:

- (i)  $R_1$  would make it morally impermissible to do anything that  $R_2$  would make it morally impermissible to do.
- (ii)  $R_1$  would make it morally impermissible do some things that  $R_2$  would *not* make it morally impermissible to do.

**Moral securitism (MS):** It is, as of  $t_i$ , *objectively morally permissible* for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively morally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . Cf. *securitism*.

**More requiring reason:** 'S has more requiring reason to  $\varphi$  than to  $\psi$  ( $\psi \neq \varphi$ )' is short for 'the set of all the reasons that S has to  $\varphi$  has greater moral requiring strength than the set of all the reasons that S has to  $\psi$  ( $\psi \neq \varphi$ )'.

**Most reason (a.k.a. optimific reason):** S has most reason to  $\varphi$  if and only if there is no *alternative*,  $\psi$ , such that S has at least as much reason to  $\psi$  as to  $\varphi$ . Cf. *optimal reason*.

**Motivating reasons for action:** This is an important subclass of *explanatory reasons for action*. These are the facts that motivated the agent to perform the act in question—that is, the facts that the agent took to be her reasons for performing the action.

**MSA:** See *maximal set of actions*. The abbreviation, 'MSAs', is short for 'maximal sets of actions'—N.B. that 'sets' is plural.

**Multiple-option case:** Multiple-option cases are cases in which there is optimal reason to achieve some end and more than one equally attractive means to achieving that end (GERT 2003, 10). An example would be a case in which I need to get to the airport and the following are all equally attractive means to my getting there: a taxi, a bus, or a train—assume that these are all equally attractive given their comparative cost, comfort, and convenience.

**Murder:** This is the act of intentionally killing an innocent human being.

**Nonconsequentialist:** 'Nonconsequentialist' is, on my usage, short for 'non-act-consequentialist'. Thus, any moral theory that is not an act-consequentialist theory is a nonconsequentialist theory, and this means that even rule-consequentialism is, in this stipulative sense, a nonconsequentialist theory.

**Non-moral reason:** This is a reason to  $\varphi$  that, morally speaking, neither counts for nor against  $\varphi$ -ing. See also *moral reason*.

**Non-requiring reason:** This is a reason that has absolutely no *moral requiring strength*. See also *requiring reason*.

**Normative reasons for action:** This is some fact that counts in favor of the agent's performing the action. Cf. *explanatory reasons for action*.

**Objectively morally impermissible:** An act is objectively morally impermissible if and only if there are decisive *moral reasons* for the agent to refrain from performing it.

**Objectively morally permissible:** An act is objectively morally permissible if and only if it is not objectively morally impermissible.

**Objectively rationally permissible:** See *objective rationality*.

**Objective ought:** S objectively ought to perform  $x$  if and only if performing  $x$  is what a normatively conscientious person would do if she faced S's choice of alternatives and had certain knowledge of all the relevant normative and non-normative facts. S objectively ought to perform some *alternative* if and only if it's the *best* alternative. Cf. *subjective ought*.

**Objective rationality:** The objective rational status of an act is purely a function of the reasons there are for and against performing it, irrespective of whether or not the agent is aware of them. Thus, an act is objectively rationally permissible if and only if it is not objectively rationally impermissible. And an act is objectively rationally impermissible if and only if there are decisive reasons for the agent to refrain from performing it. Cf. *subjective rationality*.

**Objective reason:** see *reason*.

**Obligation dilemma:** An obligation dilemma is a choice situation in which two or more of the agent's *positive act-alternatives* are obligatory. Cf. *prohibition dilemma*.

**Obligatory:** S's  $\varphi$ -ing is obligatory if and only if it is permissible for S to  $\varphi$  and impermissible for S to refrain from  $\varphi$ -ing. In other words,  $\varphi$  is obligatory if and only if  $\varphi$  is the only permissible *alternative*.

**Optimal reason:** S has optimal reason to  $\varphi$  if and only if there is no *alternative*,  $\psi$ , such that S has more reason to  $\psi$  than to  $\varphi$ . Cf. *optimific reason*.

**Optimific reason:** S has optimific reason to  $\varphi$  if and only if S has *most reason* to  $\varphi$ . Cf. *optimal reason*.

**Optional:** S's  $\varphi$ -ing is optional if and only if S is both *permitted* to  $\varphi$  and permitted to refrain from  $\varphi$ -ing.

**Outcome:** The outcome associated with S's  $\varphi$ -ing is the outcome that would obtain were S to  $\varphi$ —that is, the possible world that would be actual were S to  $\varphi$ .

**Overridingness:** To say that one type of reason, say, *m*-reasons, overrides another, say, *n*-reasons, with respect to a certain kind of normative status, *N*, is to say that, in any situation where both types of reasons are present and an act, *x*, has a certain *N*-status, no modification of the situation that involves affecting only what *n*-reasons there are will change *x*'s *N*-status. That is, if *m*-reasons override *n*-reasons with respect to an act's *N*-status, then even the weakest *m*-reason overrides the strongest *n*-reason in the determination of that act's *N*-status.

**Performs:** S *performs* a set of actions,  $\alpha_j$ , if and only if S performs each act in that set.

**Permissible:** S's  $\varphi$ -ing is permissible if and only S is not *obligated* to refrain from  $\varphi$ -ing. In other words, S's  $\varphi$ -ing is permissible if and only S's  $\varphi$ -ing is not *impermissible*.

**Personally possible:** A set of actions,  $\alpha_j$ , consisting of a number of steps extending from  $t_j$  -  $t_n$  is, as of  $t_i$ , personally possible for S if and only if S's performing the first step at  $t_j$  is, as of  $t_i$ , *securable* by S, and if S were to perform the first step at  $t_j$ , S's performing the second step at  $t_k$  would be, as of  $t_j$ , securable by S, and if S were to perform the first two steps at  $t_j$  and  $t_k$ , respectively, S's performing the third step at  $t_l$  would be, as of  $t_k$ , securable by S, and so forth until all the steps have been completed ( $t_i < t_j < t_k < t_l \dots < t_n$ ).

**Personal-value teleology:** This is the view an act's *deontic status* is determined by the personal value of its outcome, such that, if S is morally required to perform *x*, then S's performing *x* would maximize S's good. Cf. *impersonal-value teleology* and *rational-desire teleology*.

**Plan:** A plan is an intention to perform some act or set of actions in the future. For instance, one might plan to spend tomorrow morning playing golf.

**Policy:** A policy is an intention to perform a certain kind of action in certain potentially recurring situations—for example, to buckling up one's seat belt when one drives.

**Positive act:** This is the performance of an action, which contrasts with the omission of an action (i.e., a negative act).

**Possibilism:** This is the view according to which it is, as to  $t_i$ , objectively rationally permissible for S to perform a set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if the best *maximal set of actions* that is, as to  $t_i$ , personally possible for S includes S's performing  $\alpha_j$ . Cf. *actualism* and *securitism*.

**Prohibition dilemma:** A prohibition dilemma is a choice situation in which all of the agent's *positive act-alternatives* are impermissible. Cf. *obligation dilemma*.

**Proper subset:** One set of actions,  $\alpha_j$ , is a proper subset of another,  $\alpha_i$ , if and only if every element in  $\alpha_j$  is also an element in  $\alpha_i$  but not every element in  $\alpha_i$  is an element in  $\alpha_j$ —that is, if and only if  $\alpha_j \subset \alpha_i$ .

**Rational-desire teleology:** See *act-consequentialism*. Cf. *impersonal-value teleology* and *personal-value teleology*.

**Rational egoism:** This is the view that S has decisive reason to perform  $x$  if and only if, and because, S's performing  $x$  would maximize S's utility.

**Rational justifying strength:** A reason,  $R_1$ , has more rational justifying strength than another,  $R_2$ , if and only if:

- (i)  $R_1$  would make it rationally permissible to do anything that  $R_2$  would make it rationally permissible to do.
- (ii)  $R_1$  would make it rationally permissible to do some things that  $R_2$  would *not* make it rationally permissible to do.

**Rationally overriding:** Moral reasons are rationally overriding if and only if S has more moral reason to perform  $x$  than to perform  $y$  ( $y \neq x$ ) only if S is not rationally permitted to perform  $y$ . Cf. *morally overriding*.

**Rational requiring strength:** One reason,  $R_1$ , has more rational requiring strength than another,  $R_2$ , if and only if:

- (i)  $R_1$  would make it rationally impermissible to do anything that  $R_2$  would make it rationally impermissible to do.
- (ii)  $R_1$  would make it rationally impermissible to do some things that  $R_2$  would *not* make it rationally impermissible to do.

**Reason (i.e., objective reason):** There is a reason for S to  $\phi$  if and only if there is some fact that counts in favor of S's  $\phi$ -ing, where  $\phi$  is some *judgment-sensitive attitude*.

**Reasons for action:** See *normative reasons for action*.

**Requiring reason:** This is a reason that has some *moral requiring strength*. All requiring reasons are moral reasons, but it's not necessarily the case that all moral reasons are requiring reasons. There may be morally enticing reasons: reasons that, morally speaking, count in favor of  $\phi$ -ing but that don't have any moral requiring strength.

**Resolution:** This is an intention that one has resolved not to reconsider even in the face of anticipated temptation to do so. On one plausible view, a resolution involves both a first-order intention to perform a certain action and a second-order intention not to let that first-order intention be deflected by anticipated contrary inclinations—see HOLTON 2009, pp. 11-12.

**Restrictions:** See *agent-centered restrictions*.

**Revised version of moral rationalism (MR\*):** The *deontic status* of an *act-set*,  $\alpha_j$ , is determined by the agent's reasons for and against performing  $\alpha_j$ , such that, if S is, as of  $t_i$ , morally required to perform  $\alpha_j$ , then S has most reason, all things considered, to perform  $\alpha_j$ . Cf. *moral rationalism*.

**Revised version of the teleological conception of practical reasons (TCR\*):** The agent's reasons for and against performing  $\alpha_j$  are determined

by her reasons for and against preferring certain outcomes to others, such that, if S has most reason to perform  $\alpha_j$ , then, of all the outcomes that S could bring about by performing some MSA that is, as of  $t_i$ , securable by S, S has most reason to desire some subset of those that would result from S's performing an MSA that includes S's performing  $\alpha_j$ . Cf. the *teleological conception of practical reasons*.

**Rule prudence:** This is the view according to which an act is imprudent if and only if, and because, it violates the ideal code of prudential rules—i.e., the code that, if internalized by the agent, would produce, over time, more utility for her than any other available alternative code would.

**Schefflerian utilitarianism (SU):** Schefflerian utilitarianism is the view according to which all of the following hold: (1) *dual-ranking act-consequentialism* is true; (2) S has more moral reason to want  $o_i$  to obtain than to want  $o_j$  ( $j \neq i$ ) to obtain if and only if  $o_i$  contains more utility for others (i.e., for those other than S) than  $o_j$  does; and (3) S has more reason, all things considered, to want  $o_i$  to obtain than to want  $o_j$  to obtain if and only if  $o_i$  contains more egoistically-adjusted utility than  $o_j$  does, where egoistically-adjusted utility includes everyone's utility but adjusts the overall total by giving S's utility, say, ten times the weight of anyone else's.

**Securable:** A set of actions,  $\alpha_j$ , beginning at  $t_j$  is, as of  $t_i$  ( $t_i \leq t_j$ ), *securable* by S if and only if there is a set of actions,  $\alpha_i$  ( $\alpha_i \subseteq \alpha_j$ ), such that both of the following hold: (1) S would perform  $\alpha_j$  if S were, at  $t_i$ , to intend to perform  $\alpha_i$ , and (2) S is not, at  $t_i$ , in a state that's incompatible with S's intending to perform  $\alpha_i$ .

**Securitism (SEC):** The view according to which it is, as of  $t_i$ , *objectively rationally permissible* for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively rationally permissible maximal sets of actions that are, as of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . Cf. *actualism* and *possibilism*.

**Securitist consequentialism (SC):** The deontic status of a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) is determined by the reasons there are for and against the agent's preferring certain outcomes to others, such that, if S is, as of  $t_i$ , morally required to perform  $\alpha_j$ , then, of all the outcomes that S could bring about by performing some MSA that is, as of  $t_i$ , securable by S, S has most reason to desire some subset of those that would result from

S's performing an MSA that includes S's performing  $\alpha_j$ . Cf. *act-consequentialism*.

**Self-other asymmetry:** Whereas the fact that S's performing  $x$  would further S's self-interest does not constitute a moral reason for S to perform  $x$ , the fact that S's performing  $x$  would further someone else's self-interest does constitute a moral reason for S to perform  $x$ .

**Set of actions:** See *act-set*.

**Simple act:** See *minimal act*. Cf. *compound act*.

**Special obligations:** According to some moral theories, agents have certain special obligations that are specific to them as individuals given their particular relationships and history. These include obligations arising out of past acts (e.g., the obligation to keep one's promises) and also those obligations that come with occupying certain roles (e.g., professional and familial obligations). A special obligation is a type of *constraint*. There is a special obligation to perform a certain *act-type* (e.g., a special obligation to save one's own child) if and only if agents are required to perform that act-type even in some circumstances in which failing to perform that act-type is the only way to minimize comparable failures to perform that act-type.

**Subjective ought:** S subjectively ought to perform  $x$  if and only if performing  $x$  is what a normatively conscientious person would do if she were in the exact same situation that S is in, facing S's choice of alternatives and all the normative and non-normative uncertainty that goes along with being in S's epistemic position. S subjectively ought to perform some *alternative* if she believes (or, perhaps, ought to believe) that it is the *expectably best* alternative. Cf. *objective ought*.

**Subjective rationality:** On Parfit's view, an act is subjectively irrational if and only if the agent has beliefs whose truth would give her decisive reasons not to perform the act. And when the agent has inconsistent beliefs, the act will be subjectively rational relative to some beliefs but subjectively irrational relative to others (PARFIT 2008). On Gert's view, an act is subjectively irrational if and only if it indicates some failure in the practical mental functioning of the agent (GERT 2004, p. 160). (Cf. *objective rationality*.)



**Successfully counter:** To say that the reasons to  $\varphi$  successfully counter the reasons to  $\psi$  ( $\psi \neq \varphi$ ) is to say that the reasons to  $\varphi$  prevent the reasons to  $\psi$  from being *decisive* by, say, equaling, outweighing, undermining, or silencing them. Another possibility is that the reasons to  $\varphi$  are incommensurable with the reasons to  $\psi$  such that there is sufficient reason both to  $\varphi$  and to  $\psi$ .

**Sufficient reason:** S has sufficient reason to  $\varphi$  if and only if S's reasons are such as to make S *objectively rationally permitted* to  $\varphi$ . In other words, S has sufficient reason to  $\varphi$  if and only if S does not have *decisive reason* to  $\sim\varphi$ .

**Sufficient requiring reason:** A sufficient requiring reason is a reason that has sufficient *moral requiring strength* to generate, absent countervailing reasons, a moral requirement to perform the act of which it counts in favor.

**Supererogatory:** S's  $\varphi$ -ing is supererogatory if and only if, in  $\varphi$ -ing, S goes above and beyond the call of (perfect and imperfect) duty—that is, if and only if, in  $\varphi$ -ing, S does more than (perfect and imperfect) duty requires. Cf. *superperfecterogatory*.

**Superperfecterogatory:** S's  $\varphi$ -ing is superperfecterogatory if and only if, in  $\varphi$ -ing, S goes above and beyond the call of perfect duty—that is, if and only if, in  $\varphi$ -ing, S does more than perfect duty requires. Cf. *supererogatory*.

**Teleological conception of (practical) reasons (TCR):** The agent's reasons for and against performing a given act are determined by her reasons for and against preferring its outcome to those of the available alternatives, such that, if S has most (or decisive) reason to perform  $x$ , then, of all the outcomes that S could bring about, S has most (or decisive) reason to desire that  $x$ 's outcome obtains. Or to put it another way: S has more reason to perform  $a_i$  at  $t_i$  than to perform  $a_j$  at  $t_i$  just when, and because, S has more reason, at  $t_i$ , to desire that  $o_i$  ( $a_i$ 's outcome) obtains than to desire that  $o_j$  ( $a_j$ 's outcome) obtains, and S has a reason to perform  $a_i$  just when, and because, S has a reason to desire that  $o_i$  obtains. Cf. *revised version of the teleological conception of practical reasons*.

**Teleological maximizing securitism (telemax securitism or TMS):** (a) It is, as of  $t_i$ , *objectively rationally permissible* for S to perform a non-maximal set of actions,  $\alpha_j$ , beginning at  $t_j$  ( $t_i \leq t_j$ ) if and only if, and because, at least one of the objectively rationally permissible maximal sets of actions that are, as

of  $t_i$ , securable by S includes S's performing  $\alpha_j$ . (b) A maximal set of actions,  $MSA_i$ , is objectively rationally permissible if and only if, and because, its outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S. (c)  $MSA_i$ 's outcome is not, on S's evaluative ranking, outranked by that of any alternative MSA that is, as of  $t_i$ , securable by S if and only if, and because, there is no alternative MSA that is, as to  $t_i$ , securable by S whose outcome S has more reason, all things considered, to want to obtain.

**Time-identical:** Two sets of actions are time-identical if and only if they persist over the exact same interval (or intervals) of time.

**Top-down approach:** On this approach, we assess permissibility by directly applying normative principles to only maximal sets of actions and then *distributing* to determine the permissibility of non-maximal sets of actions, including those that contain only one *minimal act*. Cf. *bottom-up approach* and *across-the-board approach*.

**Traditional act-consequentialism (TAC):** This is the view according to which both (1) *act-consequentialism* is true and (2) an act produces the outcome that the agent has the most reason to desire if and only if it *maximizes* the good (impersonally construed).

**Transitive:** A relation, R, is transitive just in case: if  $xRy$  and  $yRz$ , then  $xRz$ .

**Ultimate end:** An ultimate end is "an intrinsic end that is a fundamental and indispensable part of the agent's life" (NOGGLE 2009, p. 8).

**Utile:** A utile is the smallest possible measure of *utility*—equivalent to someone's experiencing the mildest of pleasures for the briefest of moments.

**Undefeated reason:** S has an undefeated reason to  $\varphi$  if and only if S has *sufficient reason* to  $\varphi$ .

**Utilitarianism:** See *maximizing act-utilitarianism*.

**Utility:** Utility is a measure of whatever it is that enhances a subject's welfare, and disutility is a measure of whatever it is that diminishes a subject's welfare.

**Value abstractism:** This is the view that the sole or primary bearers of intrinsic value are certain abstracta—facts, outcomes, states of affairs, or possible worlds. On value abstractism, there is only one kind of value, the kind that is to be promoted, and so the only proper response to value is to desire and promote it, ensuring that there is as much of it as possible. Cf. *value concretism*.

**Value concretism:** This is the view that the fundamental bearers of intrinsic value are concrete entities (e.g., persons, animals, and things). Cf. *value abstractism*.

**Violation:** Any *infringement* of a constraint that is morally wrong is a violation of that constraint.

---

---

## Bibliography

- ALEXANDER, L. (2008). "Scalar Properties, Binary Judgments." *Journal of Applied Philosophy* 25: 85-104.
- ANDERSON, E. (1996). "Reasons, Attitudes, and Values: Replies to Sturgeon and Piper." *Ethics* 106: 538-554.
- ANDERSON, E. (1993). *Value in Ethics and Economics*. Cambridge, Mass.: Harvard University Press.
- ARNESON, R. (2002). "The End of Welfare as We Know It? Scanlon versus Welfarist Consequentialism." *Social Theory and Practice* 28: 315-336.
- ASHFORD, E. (2001). "A Response to Splawn." *Utilitas* 13: 334-341.
- BENNETT, J. (1979). "Accountability." In Z. V. Straaten (ed.), *Philosophical Subjects: Essays Presented to P. F. Strawson*, pp. 14-47. Oxford: Oxford University Press.
- BERGSTRÖM, L. (1977). "Utilitarianism and Future Mistakes." *Theoria* 43: 84-102.
- BRADLEY, B. (2006). "Against Satisficing Consequentialism." *Utilitas* 18: 97-108.
- BRATMAN, M. E. (2007). *Structures of Agency*. Oxford: Oxford University Press.
- BRATMAN, M. E. (1994). "Kagan on 'The Appeal to Cost'." *Ethics* 104: 325-332.
- BRATMAN, M. E. (1989). "Intention and Personal Policies." *Philosophical Perspectives* 3: 443-469.
- BRATMAN, M. E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, Mass.: Harvard University Press.
- BRINK, D. O. (2008). "The Significance of Desire." *Oxford Studies in Metaethics* 3: 5-46.

- BRINK, D. O. (2006). "Some Forms and Limits of Consequentialism." In D. Copp (ed.), *The Oxford Handbook of Ethical Theory*, pp. 380-423. Oxford: Oxford University Press.
- BRINK, D. O. (1997). "Kantian Rationalism: Inescapability, Authority, and Supremacy." In G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason*, pp. 255-291. New York: Oxford University Press.
- BROOK, R. (1991). "Agency and Morality." *The Journal of Philosophy* 88: 190-212.
- BROOME, J. (2004). "Reasons." In R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith, (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, pp. 28-55. Oxford: Clarendon Press.
- BROOME, J. (1991). *Weighing Goods*. Oxford: Basil Blackwell.
- BROWN, C. (2004). "Consequentialise This." Unpublished manuscript dated September 10, 2004.
- BYKVIST, K. (2009). "Objective versus Subjective Moral Oughts." Unpublished manuscript.
- BYKVIST, K. (2003). "Normative Supervenience and Consequentialism." *Utilitas* 15: 27-49.
- BYKVIST, K. (2002). "Alternative Actions and the Spirit of Consequentialism." *Philosophical Studies* 107: 45-68.
- BYRON, M. (2004). "Introduction." In his (ed.) *Satisficing and Maximizing*, pp. 1-13. Cambridge: Cambridge University Press.
- CARLSON, E. (1999). "Consequentialism, Alternatives, and Actualism," *Philosophical Studies* 96: 253-268.
- CARLSON, E. (1995). *Consequentialism Reconsidered*. Dordrecht: Kluwer Publishers.
- CASTAÑEDA, H.-N. (1989). "Moral Obligation, Circumstances, and Deontic Foci (A Rejoinder to Fred Feldman)." *Philosophical Studies* 57: 157-174.
- CHANG, R. (2004a). "Putting Together Morality and Well-Being." In P. Baumann and M. Betzler (eds.), *Practical Conflicts: New Philosophical Essays*, pp. 118-158. Cambridge: Cambridge University Press.
- CHANG, R. (2004b). "Can Desires Provide Reasons for Action?" In R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith (eds.), *Reason and Value*:

*Themes from the Moral Philosophy of Joseph Raz*, pp. 56-90. Oxford: Clarendon Press.

CHANG, R. (2004c). "All Things Considered." *Philosophical Perspectives* 18: 1-22.

CHAPPELL, R. (2009). "'Why is it that'...." (7:32 PM, November 30, 2009). Comment to a blogpost: Richard Chappell's "Actualism and Complex Actions," *Philosophy, et cetera*. Retrieved December 28, 2009, from <http://www.philosophyetc.net/2009/11/actualism-and-complex-actions.html?showComment=1259627528000#c1022235661064585180>.

CHISHOLM, R. M. (1963). "Contrary-to-Duty Imperative and Deontic Logic." *Analysis* 24: 33-36.

COCHRANE, A. (2009). "Explaining the Basic Belief." Unpublished manuscript dated March 26, 2009.

COPP, D. (1997). "The Ring of Gyges: Overridingness and the Unity of Reason." *Social Philosophy and Policy* 14: 86-106

CRISP, R. (2000). "Review of Joel Kupperman's *Value...and What Follows*." *Philosophy* 75: 458-462.

CURRAN, A. (1995). "Utilitarianism and Future Mistakes: Another Look." *Philosophical Studies* 78: 71-85.

DANCY, J. (2006). "Nonnaturalism." In D. Copp (ed.), *The Oxford Handbook of Ethical Theory*, pp. 122-145. Oxford: Oxford University Press.

DANCY, J. (2004). "Enticing Reasons." In R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, pp. 91-118. Oxford: Oxford University Press.

DARWALL, S. (2006a). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.

DARWALL, S. (2006b). "Morality and Practical Reason: A Kantian Approach." In D. Copp (ed.), *The Oxford Handbook of Ethical Theory*, pp. 282-320. Oxford: Oxford University Press.

DARWALL, S. (2002). *Welfare and Rational Care*. Princeton: Princeton University Press.

DARWALL, S. (1983). *Impartial Reason*. Ithaca: Cornell University Press.

DORSEY, D. (2009). "Weak Anti-Rationalism and the Demands of Morality." Forthcoming in *Noûs*.

- DREIER, J. (2011). "In Defense of Consequentializing." Forthcoming in *Oxford Studies in Normative Ethics*.
- DREIER, J. (2004). "Why Ethical Satisficing Makes Sense and Rational Satisficing Does Not," In M. Byron (ed.), *Satisficing and Maximizing*, pp. 131-154. Cambridge: Cambridge University Press.
- DREIER, J. (1993). "Structures of Normative Theories." *The Monist* 76: 22-40.
- ELLIS, A. (1992). "Deontology, Incommensurability, and the Arbitrary." *Philosophy & Phenomenological Research* 52: 855-875.
- EYAL, N. (2008). "Non-consequentialist Utilitarianism," *Kadish Center for Morality, Law & Public Affairs. ISUS-X, Tenth Conference of the International Society for Utilitarian Studies*. Retrieved from [http://repositories.cdlib.org/kadish/isus\\_x/by\\_N\\_Eyal](http://repositories.cdlib.org/kadish/isus_x/by_N_Eyal) on August 21, 2008.
- FELDMAN, F. (1990). "A Simpler Solution to the Paradoxes of Deontic Logic." *Philosophical Perspectives* 4: 309-341.
- FELDMAN, F. (1986). *Doing the Best We Can: An Essay in Informal Deontic Logic*. Dordrecht: D. Reidel Publishing Company.
- FERRY, M. (2009). "Does Morality Demand Our Very Best? On Moral Prescriptions and the Line of Duty." Retrieved from [http://www.colorado.edu/philosophy/center/RoME\\_2009\\_full\\_papers.shtml](http://www.colorado.edu/philosophy/center/RoME_2009_full_papers.shtml) on August 9, 2009.
- FISCHER, J. M. AND RAVIZZA, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- FOOT, P. (2001). *Natural Goodness*. Oxford, Clarendon Press.
- FOOT, P. (1985). "Utilitarianism and the Virtues." *Mind* 94: 196-209.
- GERT, J. (2009). "Moral Reasons and Rational Status." Forthcoming in *Canadian Journal of Philosophy*.
- GERT, J. (2008). "Michael Smith and the Rationality of Immoral Action." *The Journal of Ethics* 12: 1-23.
- GERT, J. (2007a). "Normative Strength and the Balance of Reasons." *Philosophical Review* 116: 533-562.
- GERT, J. (2007b). "Reply to Tenenbaum." *Canadian Journal of Philosophy* 37: 463-476.

- GERT, J. (2004). *Brute Rationality: Normativity and Human Action*. Cambridge: Cambridge University Press.
- GERT, J. (2003). "Requiring and Justifying: Two Dimensions of Normative Strength." *Erkenntnis* 59: 5-36.
- GIBBARD, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.
- GOBLE, L. (1993). "The Logic of Obligation, 'Better' and 'Worse'." *Philosophical Studies* 70: 133-163.
- GOLDMAN, H. S.—now Holly M. Smith. (1978). "Doing the Best One Can." In A. I. Goldman and J. Kim (eds.), *Values and Morals*, pp. 185-214. Dordrecht: D. Reidel Publishing Company.
- GOLDMAN, H. S.—now Holly M. Smith. (1976). "Dated Rightness and Moral Imperfection." *Philosophical Review* 85: 449-487.
- GREENSPAN, P. (1978). "Oughts and Determinism: A Response to Goldman." *Philosophical Review* 87: 77-83.
- GREENSPAN, P. (1975). "Conditional Oughts and Hypothetical Imperatives." *Journal of Philosophy* 72: 259-276
- GRIFFIN, J. (1986). *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press.
- HARMAN, E. (2009). "'I'll Be Glad I Did It' Reasoning and the Significance of Future Desires." *Philosophical Perspectives* 23: 177-199.
- HARMAN, G. (1976). "Practical Reasoning." *Review of Metaphysics* 29: 431-463.
- HARWOOD, S. (2003). "Eleven Objections to Utilitarianism." In L. Pojman (ed.), *Moral Philosophy: A Reader*, 3<sup>rd</sup> Edition, pp. 179-192. Indianapolis: Hackett Publishing.
- HEATHWOOD, C. (2008). "Fitting Attitudes and Welfare." *Oxford Studies in Metaethics* 3: 47-73.
- HILL, JR., T. E. (2002). *Human Welfare and Moral Worth*. Oxford: Oxford University Press.
- HOLTON, R. (2009). *Willing, Wanting, Waiting*. Oxford: Oxford University Press.
- HOOKE, B. (2000). *Ideal Code, Real World*. Oxford: Clarendon Press.



- HORGAN, T. and TIMMONS, M. (2010). "Untying a Knot From the Inside Out: Reflections on the 'Paradox' of Supererogation." *Social Philosophy & Policy* 27: [pp. currently unknown].
- HOWARD-SNYDER, F. (1994). "The Heart of Consequentialism." *Philosophical Studies* 76: 107-129.
- HUMBERSTONE, I. L. (1983). "The Background of Circumstances." *Pacific Philosophical Quarterly* 64: 19-34.
- HURKA, T. (1990). "Two Kinds of Satisficing." *Philosophical Studies* 59: 107-111.
- HURLEY, P. E. (2009). *Beyond Consequentialism*. Oxford: Oxford University Press.
- HURLEY, P. E. (2007). "Desire, Judgment, and Reason: Exploring the Path Not Taken." *The Journal of Ethics* 11: 436-463.
- HURLEY, P. E. (2006). "Does Consequentialism Make Too Many Demands, or None at All?" *Ethics* 116: 680-706.
- HURSTHOUSE, R. (1999). *On Virtue Ethics*. Oxford: Oxford University Press.
- JACKSON, F. and PARGETTER, R. (1986). "Oughts, Options, and Actualism." *Philosophical Review* 95: 233-355.
- JOHANSSON, J. (2009). "Fitting Attitudes, Welfare, and Time." *Ethical Theory and Moral Practice* 12: 247-256.
- JOYCE, R. (2008a). "What Neuroscience Can (and Cannot) Contribute to Metaethics." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality*, pp. 371-394. Cambridge, MA: MIT Press.
- JOYCE, R. (2008b). "Responses to Nichols and Katz." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality*, pp. 419-426. Cambridge, MA: MIT Press.
- KAGAN, S. (1994). "Defending Options." *Ethics* 104: 348-351.
- KAGAN, S. (1991). "Replies to My Critics." *Philosophy and Phenomenological Research* 51: 919-928.
- KAGAN, S. (1989). *The Limits of Morality*. Oxford: Clarendon Press.
- KAMM, F. M. (2007). *Intricate Ethics: Rights, Responsibilities, and Permissible Harms*. Oxford: Oxford University Press.

- KAMM, F. M. (1996). *Morality, Mortality: Volume II: Rights, Duties, and Status*. New York: Oxford University Press.
- KAMM, F. M. (1992). "Nonconsequentialism, the Person as an End-In-Itself and the Significance of Status." *Philosophy and Public Affairs* 21: 354-389.
- KAVKA, G. (1983). "The Toxin Puzzle." *Analysis* 43: 33-36.
- KAWALL, J. (2003). "Self-regarding Supererogatory Actions." *Journal of Social Philosophy* 34: 487-498.
- KENNETT J. (2006). "Do Psychopaths Really Threaten Moral Rationalism?" *Philosophical Explorations* 9: 69-82
- KORSGAARD, C. (1997). "The Normativity of Instrumental Reason." In G. Cullity and B. Gaut, (eds.), *Ethics and Practical Reason*, pp. 213-254. Oxford: Clarendon Press.
- LAWLOR, R. (2009a). *Shades of Goodness: Gradability, Demandingness and the Structure of Moral Theories*. New York: Palgrave Macmillan.
- LAWLOR, R. (2009b). "The Rejection of Scalar Consequentialism." *Utilitas* 21: 100-116.
- LEVY, N. (2005). "The Good, the Bad and the Blameworthy." *Journal of Ethics & Social Philosophy*, [www.jesp.org](http://www.jesp.org), vol. 1, no. 2.
- LOPEZ, T., ZAMZOW, J., GILL, M., and NICHOLS, S. (2009). "Side Constraints and the Structure of Commonsense Ethics." *Philosophical Perspectives* 23: 305-319.
- LORD, E. (2009). "Having Reasons and the Factoring Account." Forthcoming in *Philosophical Studies*.
- LOUISE, J. (2008). "I Won't Do It! Self-Prediction, Moral Obligation and Moral Deliberation." *Philosophical Studies* DOI: 10.1007/s11098-008-9258-5 (published online August 7, 2008).
- LOUISE, J. (2004). "Relativity of Value and the Consequentialist Umbrella," *Philosophical Quarterly* 54: 518-536.
- MCCONNELL, T. C. (1980). "Utilitarianism and Supererogatory Acts." *Ratio* 22: 36-38.
- MCLEOD, O. (2001). "Just Plain 'Ought'." *The Journal of Ethics* 5: 269-291.
- MCMAMARA, P. (2011). "Action Beyond Morality's Call Versus Supererogatory Action: Toward a More Adequate Conceptual Scheme

- for Common Sense Morality." Forthcoming in *Oxford Studies in Normative Ethics*.
- MCNAMARA, P. (1996a). "Making Room for Going Beyond the Call." *Mind* 105: 415-450.
- MCNAMARA, P. (1996b). "Must I Do What I Ought (or Will the Least I Can Do Do)?" In M. A. Brown and J. Carmo (eds.), *Deontic Logic, Agency and Normative Systems*, pp. 154-173. Berlin: Springer-Verlag.
- MCNAUGHTON, D. AND RAWLING, P. (1991). "Agent-Relativity and the Doing-Happening Distinction." *Philosophical Studies* 63: 167-185.
- MELLEMA, G. (1991). "Supererogation and the Fulfillment of Duty." *The Journal of Value Inquiry* 25: 167-175.
- MILL, J. S. (1991). [1861]. *Utilitarianism*. In J. M. Robson (ed.), *Collected Works of John Stuart Mill*, vol. 10, pp. 203-59. London: Routledge.
- MOLLER, D. (2009). "Review of Michael Zimmerman's *Living with Uncertainty: The Moral Significance of Ignorance*." *Ethics* 119: 606-611.
- MOORE, E. (2007). "Objective Consequentialism, Right Actions, and Good People." *Philosophical Studies* 133: 83-94.
- NAGEL, T. (1978). *The Possibility of Altruism*. Princeton: Princeton University Press.
- NICHOLS, S. (2008). "Moral Rationalism and Empirical Immunity." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality*, pp. 395-408. Cambridge, MA: MIT Press.
- NICHOLS, S. (2002). "How Psychopaths Threaten Moral Rationalism, Or Is It Irrational to be Amoral?" *The Monist* 85: 285-304.
- NOGGLE, R. (2009). "Give Till It Hurts? Beneficence, Imperfect Duties, and a Moderate Response to the Aid Question." *Journal of Social Philosophy* 40: 1-16.
- NORCROSS, A. (2006a). "The Scalar Approach to Utilitarianism." In H. West (ed.), *The Blackwell Guide to Mill's Utilitarianism*, pp. 217-232. Oxford: Oxford University Press.
- NORCROSS, A. (2006b). "Reasons without Demands: Rethinking Rightness." In J. Dreier (ed.), *Contemporary Debates in Moral Theory*, pp. 5-20. Oxford: Blackwell.

- ODDIE, G. & MILNE, P. (1991). "Act and Value: Expectation and the Representability of Moral Theories." *Theoria* 57: 42-76.
- OLSON, J. (2009). "Fitting Attitude Analyses of Value and the Partiality Challenge." *Ethical Theory and Moral Practice* DOI: 10.1007/s10677-009-9176-4 (published online May 5, 2009).
- PARFIT, D. (2008). *On What Matters*. Unpublished manuscript dated September 28, 2008.
- PARFIT, D. (2001). "Rationality and Reasons." In D. Egonsson, J. Josefsson, B. Petersson, and T. Rønnow-Rasmussen (eds.), *Exploring Practical Philosophy*, pp. 17-39. Aldershote: Ashgate.
- PARFIT, D. (1997). "Reasons and Motivation." *Proceeding of the Aristotelian Society*, Supplementary Volume 71: 99-130.
- PILLER, C. (2006). "Content-Related and Attitude-Related Reasons for Preferences." *Royal Institute of Philosophy Supplement* 81: 155-182.
- PORTMORE, D. W. (2011). "Consequentialism and Moral Rationalism." Forthcoming in *Oxford Studies in Normative Ethics*.
- PORTMORE, D. W. (2010). "Imperfect Reasons and Rational Options." Forthcoming in *Noûs*.
- PORTMORE, D. W. (2009a). "The Teleological Conception of Practical Reasons." Unpublished manuscript dated April 21, 2009.
- PORTMORE, D. W. (2009b). "Consequentializing." *Philosophy Compass* 4: 329-347.
- PORTMORE, D. W. (2008a). "Are Moral Reasons Morally Overriding?" *Ethical Theory and Moral Practice* 11: 369-388.
- PORTMORE, D. W. (2008b). "Dual-Ranking Act-Consequentialism." *Philosophical Studies* 138: 409-427.
- PORTMORE, D. W. (2007a). "Consequentializing Moral Theories." *Pacific Philosophical Quarterly* 88: 39-73.
- PORTMORE, D. W. (2007b). "Welfare, Achievement, and Self-Sacrifice." *Journal of Ethics & Social Philosophy*, [www.jesp.org](http://www.jesp.org), vol. 2, no. 2.
- PORTMORE, D. W. (2005). "Combining Teleological Ethics with Evaluator Relativism: A Promising Result." *The Pacific Philosophical Quarterly* 86: 95-113.

- PORTMORE, D. W. (2003). "Position-Relative Consequentialism, Agent-Centered Options, and Supererogation." *Ethics* 113: 303-332.
- PORTMORE, D. W. (2001). "Can an Act-Consequentialist Theory be Agent-Relative?" *American Philosophical Quarterly* 38: 363-77.
- PORTMORE, D. W. (2000). "Commonsense Morality and Not Being Required to Maximize the Overall Good." *Philosophical Studies* 100: 193-213.
- PORTMORE, D. W. (1998). "Can Consequentialism Be Reconciled with Our Common-Sense Moral Intuitions?" *Philosophical Studies* 91: 1-19.
- QUINN, W. S. (1993). "Putting Rationality in its Place." In his *Morality and Action*. Cambridge: Cambridge University Press.
- QUINN, W. S. (1990). "The Puzzle of the Self-Torturer." *Philosophical Studies* 59: 70-90.
- RABINOWICZ, W. and RØNNOW-RASMUSSEN, T. (2004). "The Strike of the Demon: On Fitting Pro-Attitudes and Value." *Ethics* 114: 391-423.
- RAWLS, J. (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- RAZ, J. (1999). *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press.
- RAZ, J. (1975). "Permissions and Supererogation." *American Philosophical Quarterly* 12: 161-168.
- RIDGE, M. (2005). "Review of Pleasure and the Good Life." *Mind* 114: 414-417.
- ROBERTSON, S. (2009). "Reasons, Values and Morality." In J. Skorupski (ed.), *The Routledge Companion to Ethics*. London: Routledge.
- RØNNOW-RASMUSSEN, T. (2009). "On For Someone's Sake Attitudes." *Ethical Theory and Moral Practice* 12: 397-411.
- RØNNOW-RASMUSSEN, T. (2007). "Analysing Personal Value." *Journal of Ethics* 11: 405-435.
- ROSS, J. (2009). "How to be a Cognitivist about Practical Reason." *Oxford Studies in Metaethics* 4: 242-282.
- ROSS, W. D. (1930). *The Right and the Good*. Oxford: Oxford University Press.
- SACHS, B. (2009). "Consequentialism's Double-Edged Sword." Forthcoming in *Utilitas*.

- SCANLON, T. M. (2007). "Structural Irrationality." In G. Brennan, R. Goodin, F. Jackson, and M. Smith (eds.), *Common Minds: Themes from the Philosophy of Philip Pettit*, pp. 84-103. Oxford: Oxford University Press.
- SCANLON, T. M. (2001). "Sen and Consequentialism." *Economics and Philosophy* 17: 39-50.
- SCANLON, T. M. (1998). *What We Owe to Each Other*. Cambridge, Mass.: The Belknap Press.
- SCHEFFLER, S. (1992). *Human Morality*. New York: Oxford University Press.
- SCHEFFLER, S. (1994). *The Rejection of Consequentialism*. Revised edition. Oxford: Clarendon Press. The original addition was published in 1982.
- SCHROEDER, M. (2009). "Means-End Coherence, Stringency, and Subjective Reasons." *Philosophical Studies* 143: 223-248.
- SCHROEDER, M. (2008a). "Value and the Right Kind of Reason." Unpublished manuscript dated August 10, 2008.
- SCHROEDER, M. (2008b). "Value Theory." *The Stanford Encyclopedia of Philosophy (Spring 2008 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2008/entries/value-theory/>>.
- SCHROEDER, M. (2008c). "Having Reasons." *Philosophical Studies* 139: 57-71.
- SCHROEDER, M. (2007a). "Teleology, Agent-Relative Value, and 'Good'." *Ethics* 117: 265-295.
- SCHROEDER, M. (2007b). *Slaves of the Passions*. Oxford: Oxford University Press.
- SEPIELLI, A. (2010). 'Along an Imperfectly-Lighted Path': *Practical Rationality and Normative Uncertainty*. Ph.D. dissertation, Rutgers University, United States – New Jersey. Retrieved [??], 2010, from Dissertations & Theses: Full Text. (Publication No. AAT [??]).
- SEPIELLI, A. (2009). "What to Do When You Don't Know What to Do." *Oxford Studies in Metaethics* 4: 5-28.
- SHAFFER-LANDA, R. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.
- SIDER, T. (1993). "Asymmetry and Self-Sacrifice." *Philosophical Studies* 70: 117-132.
- SIDGWICK, H. (1966). [1907]. *The Methods of Ethics*. New York: Dover.

- SINGER, P. (1999). "A Response." In D. Jamieson (ed.), *Singer and His Critics*, pp. 269-335. Oxford: Blackwell.
- SINNOTT-ARMSTRONG, W. (2005). "You Ought to be Ashamed of Yourself (When You Violate an Imperfect Moral Obligation)." *Philosophical Issues* 15: 193-208.
- SINNOTT-ARMSTRONG, W. (2003). "Consequentialism." *The Stanford Encyclopedia of Philosophy (Summer 2003 Edition)*, Edward N. Zalta (ed.), URL=<<http://plato.stanford.edu/archives/sum2003/entries/consequentialism/>>.
- SKORUPSKI, J. (2009). *The Domain of Reasons*. Unpublished manuscript dated January 20, 2009.
- SKORUPSKI, J. (1999). *Ethical Explorations*. Oxford: Oxford University Press.
- SLOTE, M. (2001). *Morals from Motives*. Oxford: Oxford University Press.
- SLOTE, M. (1991). "Shelly Kagan's *The Limits of Morality*." *Philosophy and Phenomenological Research* 51: 915-917.
- SLOTE, M. (1989). *Beyond Optimizing: A Study of Rational Choice*. Cambridge, Mass.: Harvard University Press.
- SLOTE, M. (1985). *Common-Sense Morality and Consequentialism*. London: Routledge and Kegan Paul.
- SLOTE, M. (1984a). "Satisficing Consequentialism," *Proceedings of the Aristotelian Society* 58: 139-63.
- SLOTE, M. (1984b). "Morality and Self-Other Asymmetry," *Journal of Philosophy* 81: 179-192.
- SMITH, A. M. (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115: 236-271.
- SMITH, A. M. (2008). "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 138: 367-392.
- SMITH, H. M. (2011). "The Moral Clout of Reasonable Beliefs." Forthcoming in *Oxford Studies in Normative Ethics*.
- SMITH, M. (2009). "Two Kinds of Consequentialism." Forthcoming in *Philosophical Issues*.
- SMITH, M. (1994). *The Moral Problem*. Oxford: Blackwell.

- SOBEL, D. (2009). "Subjectivism and Blame." Forthcoming in the *Canadian Journal of Philosophy*.
- SOBEL, D. (2007). "The Impotence of the Demandingness Objection." *Philosophers' Imprint* 7: 1-17.
- SOBEL, J. H. (1994). *Taking Chances: Essays on Rational Choice*. Cambridge: Cambridge University Press.
- SOBEL, J. H. (1976). "Utilitarianism and Past and Future Mistakes." *Noûs* 10: 195-219.
- SOSA, D. (1993). "Consequences of Consequentialism." *Mind* 102: 101-22.
- SPLAWN, C. (2001). "The Self-Other Asymmetry and Act Utilitarianism." *Utilitas* 13: 323-333.
- STOCKER, M. (1967). "Acts, Perfect Duties, and Imperfect Duties." *Review of Metaphysics* 20: 507-517.
- STRATTON-LAKE, P. and HOOKER, B. (2006). "Scanlon versus Moore on Goodness." In T. Horgan and M. Timmons (eds.), *Metaethics after Moore*, pp. 149-168. Oxford: Clarendon Press.
- STRATTON-LAKE, P. (2005). "How to Deal with Evil Demons: Comment on Rabinowicz and Rønnow-Rasmussen." *Ethics* 115: 788-798.
- STRAWSON, P. (1962). "Freedom and Resentment." *Proceedings of the British Academy* 48: 187-211.
- STROUD, S. (1998). "Moral Overridingness and Moral Theory." *Pacific Philosophical Quarterly* 79: 170-189.
- STURGEON, N. L. (1996). "Anderson on Reason and Value." *Ethics* 106: 509-524.
- SUIKKANEN, J. (2009a). "Consequentialist Options." Retrieved from [http://www.colorado.edu/philosophy/center/RoME\\_2009\\_full\\_papers.shtml](http://www.colorado.edu/philosophy/center/RoME_2009_full_papers.shtml) on August 9, 2009.
- SUIKKANEN, J. (2009b). "Consequentialism, Constraints and the Good-Relative-to: A Reply to Mark Schroeder." *Journal of Ethics & Social Philosophy*, [www.jesp.org](http://www.jesp.org), March 2009.
- SUMNER, L. W. (1996). *Welfare, Happiness, and Ethics*. Oxford: Oxford University Press.
- TÄNNSJÖ, T. (1999). "A Concrete View of Intrinsic Value." *The Journal of Value Inquiry* 33: 531-536.



- TAUREK, J. (1977). "Should the Numbers Count?" *Philosophy and Public Affairs* 6: 293-316.
- TENENBAUM, S. (2007). "Brute Requirements: A Critical Notice of J. Gert's *Brute Rationality*." *Canadian Journal of Philosophy* 37: 153-172.
- THOMASON, R. H. (1981). "Deontic Logic and the Role of Freedom in Moral Deliberation." In R. Hilpinen (ed.), *New studies in deontic logic*, pp. 177-186. Dordrecht: D. Reidel Publishing Company.
- THOMSON, J. J. (1990). *The Realm of Rights*. Cambridge, Mass.: Harvard University Press.
- UNGER, P. (1996). *Living High and Letting Die*. Oxford: Oxford University Press.
- VAN ROOJEN, M. (2009). "Moral Rationalism and Rational Amoralism." Unpublished manuscript dated February 2009.
- VAN ROOJEN, M. (2004). "The Plausibility of Satisficing and the Role of Good in Ordinary Thought." In M. Byron (ed.), *Satisficing and Maximizing*, pp. 1-13. Cambridge: Cambridge University Press.
- VALLENTYNE, P. (1989). "Two Types of Moral Dilemmas." *Erkenntnis* 30: 301-318.
- VALLENTYNE, P. (1988). "Gimmicky Representation of Moral Theories." *Metaphilosophy* 19: 253-63.
- VELLEMAN, J. D. (1993). "Well-Being and Time." Reprinted in John Martin Fischer (ed.), *The Metaphysics of Death*, pp. 329-3. Stanford: Stanford University Press. The article originally appeared in *Pacific Philosophical Quarterly* 72 (1991): 48-77.
- VELLEMAN, J. D. (1993). *Practical Reflection*. Princeton: Princeton University Press.
- VESSEL, J.-P. (2010). "Supererogation for Utilitarianism." Forthcoming in the *American Philosophical Quarterly*.
- VESSEL, J.-P. (2009). "Defending a Possibilist Insight in Consequentialist Thought." *Philosophical Studies* 142: 183-195.
- WALLACE, R. J. (2005). "Reasons, Values, and Agent-Relativity." Unpublished paper dated June 8, 2005.

- WALLACE, R. J. (2003). "Practical Reason." *The Stanford Encyclopedia of Philosophy (Winter 2003 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2003/entries/practical-reason/>.
- WALLACE, R. J. (2001). "Normativity, Commitment, and Instrumental Reason." *Philosophers' Imprint* 1.
- WALLACE, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press.
- WEDGWOOD, R. (2009). "Against Actualism." *PEA Soup*, URL = <http://peasoup.typepad.com/peasoup/2009/09/against-actualism.html>
- WILLIAMS, B. (1973). "A Critique of Utilitarianism." In J. J. C. Smart and B. Williams (eds.), *Utilitarianism: For and Against*, pp. 75-150. Cambridge: Cambridge University Press.
- WOLF, S. (1990). *Freedom within Reason*. New York. Oxford University Press.
- WOLF, S. (1986). "Above and Below the Line of Duty," *Philosophical Topics* 14: 131-48
- WOODARD, C. (2009). "What's Wrong with Possibilism?" *Analysis* 69: 219-226.
- ZIMMERMAN, M. J. (2009). "Partiality and Intrinsic Value." Unpublished manuscript.
- ZIMMERMAN, M. J. (2008). *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.
- ZIMMERMAN, M. J. (2006). "The Relevance of Risk to Wrongdoing." In K. McDaniel, J. R. Raibley, R. Feldman, and M. J. Zimmerman (eds.), *The Good, the Right, Life and Death: Essays in Honor of Fred Feldman*, pp. 151-170. Burlington, VT: Ashgate Publishing Company.
- ZIMMERMAN, M. J. (1996). *The Concept of Moral Obligation*. Cambridge: Cambridge University Press.
- ZIMMERMAN, M. J. (1993). "Supererogation and Doing the Best One Can." *American Philosophical Quarterly* 30: 373-380.