

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Philosophy

2013

Values, Reasons, & Ought

Maguire, Barry

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp0105741r83d>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

Values, Reasons, & Ought

Barry Maguire

**A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY**

**RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PHILOSOPHY**

[Advisor: Gideon Rosen]

September 2013

© Copyright by Barry Maguire, 2013. All rights reserved.

ABSTRACT

This dissertation presents a systematic account of the explanatory relations between our most basic evaluative and normative notions: *values*, *reasons*, and *ought*. I defend three central theses. The first thesis, the value-based theory of reasons, maintains that for some fact to be a reason for you to X is for that fact to constitute or explain the fact that your X-ing would realise some state of affairs that is valuable. The fact that giving flowers to Alfred will make him happy is a reason to do so, since his happiness is valuable. The second thesis concerns the weight of reasons. The weight of any reason is an increasing function of the amount of value likely to be realised, and the value of your having a pro-attitude towards that value. In this case the value of the pro-attitude will be explained by the nature of your relationship with Alfred. Importantly, while values are neutral, weighted reasons are always relative. The positive account vindicates deontological intuitions about partiality, integrity, and commitment, and virtue-theoretic intuitions about the value and normative significance of attitudes and the live well-lived. The third thesis, the reasons-based theory of ought, maintains that for some X to be what you ought to do, all things considered, is for the net weight of the reasons in favour of X to be greater than the net weight of the reasons in favour of any alternative to X. So if your reasons to give flowers to Alfred are weightier than your reasons to go apple-picking with Mildred, and weightier than your reasons to do anything else, then you ought to give flowers to Alfred.

ACKNOWLEDGEMENTS

My greatest thanks go to three people: Gideon Rosen, Jack Woods, and Meghan Flaherty. Gideon is a tremendous teacher. His modest questions here and there invariably led to complete reworkings of my ideas. The entire second chapter fell out of one remark of his, which I'm sure he doesn't remember, that a certain consequentialist account of partiality (to which I was attracted at the time) 'seems to get the right results for the wrong reasons.' His own philosophical thinking is rigorous, penetrating, and clear, and always characterised by a deep sense of what is important about a problem, and what is important about the presentation of a putative solution. He has had to work hard to get me to manifest any of these virtues to any degree in my own writing, and I am extremely grateful to him for his efforts. Jack Woods is as much of a mentor. The poor guy has read everything I have written in the last several years, usually in every draft, and dutifully supplied counterexamples on every occasion. I owe a great deal to his philosophical breadth and passion, and to his generosity of spirit. To Meghan Flaherty go my greatest thanks. She inspires me every day with her love, her talent, and her motivation.

I have been extremely fortunate to study in two great intellectual environments, in each of which I have learned by rubbing shoulders with some wonderful colleagues and teachers. From Princeton, I would especially like to thank Michael Smith, from whom I have learned a huge amount. His graduate seminars were always clear and challenging, and our conversations over the years have changed what I think about many issues. I would also like to thank Anthony Appiah, Derek Baker, Arudra

Burra, Ryan Cook, Richard Yetter Chappell, Anthony Cross, Simon Cullen, Nate Gadd, Elizabeth Harman, Gilbert Harman, Daniel Herrick, Sukaina Hirji, Frank Jackson, Mark Johnston, Boris Kment, Raphael Krut-Landau, Adam Lerner, Sarah-Jane Leslie, Eden Lin, Errol Lord, Hendrik Lorenz, Tristram McPherson, Carla Merino, Philip Pettit, Kristin Primus, Joseph Rachiele, Ryan Robinson, Whitney Schwab, Peter Singer, Jack Spencer, Noel Swanson, Cecilia Tilli, Gerard Vong, and Daniel Wodak. From Oxford, I'd especially like to thank Adam Swift, a really wonderful teacher, and the sadly late Jerry Cohen. I'd also like to thank Carlos Ardila, Tom Dougherty, the sadly late Bob Hargrave, Markus Kneer, George Mitton, Jaideep Nandwani, Mary Robertson, Paulina Sliwa, and Rowland Stout.

From the profession, I'd like to thank Ralf Bader, for an inspiring seminar on the metaphysics of value. Jonathan Dancy, for reading and thinking about chapter three, and complaining only a little. Fred Feldman, for telling me that axiologists don't take buck-passing seriously. David Frank for help with the parts of chapter three concerning contributory value. Warren Goldfarb, for wonderful conversations about Wittgenstein. I'm afraid Professor Goldfarb would not approve of the unabashed theory-building within these pages, but perhaps I can return to his lessons now that the dissertation is out of the way. Antti Kauppinen, for discussions about the value-based theory of reasons and the relation between ought and criticism. Stephen Kearns, for good craic and for encouraging the paper on supererogation. Geoff Sayre-McCord, for very generously taking the time to think and talk through the earliest version of chapter two. Tim Scanlon, for several helpful conversations about reasons fundamentalism. Karl Schafer, for extremely helpful comments on various drafts of two chapters. Nick Stang, for being himself.

Daniel Star, for helpful discussions about value-based and evidence-based theories of reasons. Kurt Sylvan, for advice about value-based theories of epistemic reasons. An audience at Georgia State, for helpful discussions of the early material on particularism and value-based theories. And an audience at the Northern Institute of Philosophy for helpful discussions about grounding and explanation.

From close family and friends I'd like to thank Scott Beveridge, my greatest friend, for solidarity and for jokes in the mornings. Jack Woods, again. Paul Forsyth, a great man and a great friend. Gerry O'Hare, Richard Spence, and Paul Graham, for many years of friendship and encouragement, even after I started going to bed early. Jan Blackstone, for my earliest philosophical conversations. Cory Wright, for companionship and support during the last days. Tom and Eileen, for constant support throughout the years. My Dad, for inspiration and encouragement. Laura and AJ, and now Dru, for being wonderful siblings and great friends.

I dedicate this dissertation to my Mum, Mary Maguire, a really amazing woman, a great parent and friend, and someone whose unyieldingly selfless love and concern, both for me and for everyone else she knows, continues to astound me. I am inexpressibly grateful to her.

TABLE OF CONTENTS

Abstract	iii
Acknowledgements	iv
Introduction	ix
Chapter One: Values and Reasons	1
Abstract	1
Introduction	1
Section One: The Value-Based Theory of Reasons	3
Section Two: The Buck-Passing Theory of Value	9
Scanlon's First Argument: From Explanation	12
Scanlon's Second Argument: From Pluralism	14
Section Three: Arguments for the Value-Based Theory of Reasons	16
Argument One: Structural Advantages	16
Argument Two: Counting Reasons	22
Argument Three: Weighing Reasons	27
Conclusion	35
Chapter Two: A Neutral Value-Based Theory of Partial Reasons	37
Abstract	37
Section One: The Problem	38
Section Two: The Positive Proposal	44

(1) The relevant attitudes are valuable.	46
(2) Why is it good to have partial attitudes?	49
(3) The weights of reasons covary with the values of the relevant attitudes.	53
Why the Harmony Principle Works	56
Section Three: The Theoretical Alternatives	59
Reasons Consequentialism	59
Global Consequentialism	60
Indirect Consequentialism	63
Conclusion	67
 Chapter Three: Ought, Most Reason, & Supererogation	 70
Abstract	70
Analysing Ought in terms of Most Reason	71
A Schema for Putative Counterexamples	74
Matters Ideological	78
Section One: Requirements	80
Section Two: Permissibility	86
Singerian Cases	89
Rights and Entitlements	90
Imperfect Duties	92
Section Three: Criticism	99
Conclusion	103
 Works Cited	 104

INTRODUCTION

My Dad is a default deontologist. When I was growing up, the rules of the house were thus and so, and no sense could be made of the request for an explanation of them, either in general or in particular cases. One central recollection from my years as a recalcitrant youth is a deep frustration with this way of thinking – or not thinking – about ethics. Of course, at some point any parent's spade is turned, and any ethicist's. But there are explanations and explanations: some have more depth, more structure, more intuitiveness than others. Several years later, I was working on particularism, attempting to find some pattern in the exceptions to familiar moral principles. I became increasingly attracted by the idea that we should think of principles as heuristics that direct us towards or away from action-types that tend to realise states that are valuable or disvaluable. Plausibly there is no reason to keep a promise to do great harm; plausibly there is no reason not to pursue the pain constitutive of athletic improvement; plausibly there is no reason not to engage in sexual acts which will not lead to physical or psychological harm or violations of trust. *Ceteris paribus* moral principles are straightforwardly defeated in cases in which the relevant value or disvalue would not be realised. Alas, there is plenty of Dancy-an complexity in the metaphysics of value, so this thought does not quite solve the problems in the holism and particularism literature. But the central idea – that normative matters are to be explained by appealing to facts about value – strikes me as correct. It is an old idea. "Every art and every inquiry, and similarly

every action and every pursuit, is thought to aim at some good; and for this reason the good has rightly been declared to be that at which all things aim.”¹ It is often simply assumed in debates taking place in pubs next to politics and practical ethics departments. The fundamental normative significance of values drives G.A. Cohen’s influential critique of John Rawls, for instance. If our spade must be turned, the best place is to turn it is at the fact that this or that is valuable. It is most plausibly here that no more can be said.

So we seek to explain normative facts in terms of facts about value. But which kinds of normative facts? In recent history, W.D. Ross did the most to direct the attention of normative theorists to *pro tanto* considerations – considerations which *count in favour* of this or that. He did so, roughly speaking, by pointing out an ambiguity in claims about duties. Claims about duties are systematically ambiguous between all-things-considered claims about requirements, or about what one ought to do in a situation (our ‘duty proper’), and more general *pro tanto* claims, for instance about obligations (‘prima facie duties’). I have a duty to visit my grandmother, and to give money to charity, but it is not the case that I ought to do these things right now. Here the literature on particularism is helpful again. The fact is that there are very few, if any, interesting principles relating action-types with the all-things-considered ought. Take any action you like, no matter how gruesome, and a sufficiently imaginative ethicist will describe you a situation in which, alas, that is the thing you ought to do. It is natural for a theorist to look instead at these considerations which count in favour of certain actions, and try to give some account of them. Since Kurt Baier’s *The Moral Point of View* in 1958 it has been common to theorise about these

¹ This is the opening sentence of the *Nicomachean Ethics*.

pro tanto considerations, or *reasons*; this trend increased exponentially after the 1970 publication of Thomas Nagel's *The Possibility of Altruism*.

An under-developed part of the theory of reasons concerns their weight, and how facts about weight are explained in terms of other facts in the situation. Mark Schroeder taught us that we can distinguish the explanation of the existence of a reason from the explanation of the weight of that reason. This key insight gives us the theoretical space to offer a more nuanced account of the weight of reasons. These facts about weights in turn generate facts about what there is *most* reason to do. And in turn, the facts about what there is most reason to do generate facts about what one ought to do, all things considered. So, at least, I shall argue.

This dissertation takes the first steps towards the full defence of an account of the structure of normativity, according to which facts about value explain facts about reasons and their weights, which in turn explain facts about ought. Here follows abstracts for the three chapters.

CHAPTER ONE: VALUES AND REASONS

I present and argue for the value-based theory of reasons. This is the view that for some fact to be a reason for you to do something is for that fact to either constitute or partly explain the fact that your doing that thing realises some state of affairs that is valuable. I compare the value-based theory of reasons with an influential

alternative, T.M. Scanlon's buck-passing theory of value. I argue that the value-based theory of reasons has structural advantages over the buck-passing theory, and that it enables us to count reasons, and to give a better account of their weight.

CHAPTER TWO: A NEUTRAL VALUE-BASED THEORY OF PARTIAL REASONS

There are several powerful motivations for value-based theories of normativity, of which the various forms of consequentialism are the most famous instances. However value-based theories of normativity, in the form of these most famous instances, encounter serious difficulties accounting for a number of pervasive and important ethical phenomena. These phenomena all involve partiality of one form or another; specifically they concern one's personal relationships, personal projects, and ethical and voluntary commitments. I argue that these difficulties result from an overly simplistic conception of the resources available to value-based theories of normativity, and specifically from an overly simplistic conception of the relation between various kinds of facts about value and facts about the weights of reasons. I present and defend an alternative ethical principle which not only accommodates our intuitions about partiality, but fully vindicates these intuitions. According to this principle, the weight of a reason for you to do something is an increasing function of the amount of value that explains the reason in question and the value of your having a pro-attitude towards this value. I argue that this principle preserves what is distinctively attractive about prominent forms of consequentialism (act, global, and indirect consequentialism), while avoiding their distinctive

shortcomings. The principle also accommodates attractive features of deontology and virtue ethics.

CHAPTER THREE: OUGHT, MOST REASON, AND SUPEREROGATION

An increasing number of contemporary ethical theorists are attracted to a reasons-first account of normativity, and hence to a reasons-based analysis of the all-things-considered ought. The most natural such analysis maintains that one ought to do whatever there is most reason to do. However there is an important class of putative counterexamples to this analysis, which can all loosely be described as supererogatory actions. In such cases it appears that there are weighty reasons to do something, and in some cases these reasons seem weightier than the reasons favouring any alternatives, but it is not the case that one ought to do that thing. I argue that none of these cases constitutes a counterexample to the analysis of ought in terms of most reason.

There are two classes of cases to consider. One class involves alternatives with some positive deontic status: one is required, obligated, duty-bound to do the non-supererogatory thing. I argue that in such cases we need to weaken the principles relating these deontic notions to the all things considered ought. Sometimes you ought to do the thing that is required, and sometimes not. Another class of cases involves alternatives which are all putatively permissible. In some of these cases the alternatives will be equally well supported by reasons; such cases pose no obstacle

to our analysis. I argue that in others, we really ought to do whatever is favoured by the balance of reasons. I present two considerations in support of this claim: a no-relevant-difference argument, and some remarks about the relation between ought and the appropriateness of criticism.

The most interesting cases of this latter sort involve ‘imperfect duties.’ It is characteristic of imperfect duties that they do not require that one do anything *at a time*. Rather they require that one do *enough* of a certain kind of action over a certain period of time. I appeal to a notion from contemporary axiology, that of *contributory value*, arguing that we can use this notion to explain the reasons and their weights in such a way as to get the right results in these cases, without troubling the analysis of ought.

CHAPTER ONE: VALUES AND REASONS

ABSTRACT

I present and argue for the value-based theory of reasons. This is the view that for some fact to be a reason for you to do something is for that fact to either constitute or partly explain the fact that your doing that thing realises some state of affairs that is valuable. I compare the value-based theory of reasons with an influential alternative, T.M. Scanlon's buck-passing theory of value. I argue that the value-based theory of reasons has structural advantages over the buck-passing theory, and that it enables us to count reasons, and to give a better account of their weight.

INTRODUCTION

What is the relation between values and reasons? I hold before me a 15th century Ming vase. I have reasons to treat this vase with care. The vase is very valuable. Is the vase valuable because I have reasons to treat it with care, or have I reasons to treat it with care because it is valuable? Dropping a bomb on

Nagasaki would destroy thousands of innocent lives. This would be terrible, and it constitutes a reason not to do so. But which explains which?

The traditional view handed down from Plato and Aristotle has it that facts about values and disvalues explain the facts about normativity. In more recent times this direction of explanation has been controverted. An illustrious string of normative theorists, including Henry Sidgwick, Franz Brentano, A.C. Ewing, Paul Ziff, and most significantly in recent times, T.M. Scanlon and Jonathan Way, have attempted instead to understand value in terms of reasons – reasons to have certain desires or emotions, or reasons to respond to objects or states of affairs in certain positive ways.¹

I will resist this, and make the case for the value-based theory of reasons. In the loud and vigorous debates about reasons in recent literature, this position has been curiously overlooked. Only one or two prominent philosophers have defended a similar position.² Most value-first theorists of normativity have focused on providing a theory of *ought*, and have become consequentialists. In my view this is a mistake. Moreover this is a mistake that has prompted a misunderstanding of the strength of value-based theories of normativity, which in turn has made alternative theories seem more attractive than they should.

¹ Sidgwick, H., 1874; Brentano, F, 1889; Ewing, A.C. 1939, 1947; Ziff, P. 1960; Scanlon, T.M. 1998, Way, J. 2013. For a little more history and a list of references, see Olson, J. (forthcoming).

² In recent times Joseph Raz is the most famous exponent of something like a value-based theory of reasons; see his 1999 and 2012; for reasons to doubt that this is his view, see his 2013. Ralph Wedgwood has begun to develop a similar view in his 2009a, 2009b and *forthcoming*. Then we have various other snippets and suggestions, mostly in the context of objecting to other views: in Pekka Väyrynen 2006; in G.A. Cohen 2008; in Andrew Reisner 2009; in John Brunero *forthcoming*; and there is a sketch of a similar view in Eric Wiland's 2012.

As we will see, a value-first normative theorist can enjoy a great many theoretical fruits by focussing her attentions on *reasons* rather than *ought*.

This paper has three sections. Section One introduces the value-based theory of reasons. Section Two discusses the buck-passing view and Scanlon's two arguments for the view. I argue that Scanlon's first argument does not support the buck-passing view. His second argument presents a challenge: that the appeal to value is otiose. Section Three takes up this challenge, arguing that the appeal to value does important explanatory work.

SECTION ONE: THE VALUE-BASED THEORY OF REASONS

A reason is a fact that favours, or counts in favour of, or speaks in favour of, something or other. The fact that the bakery sells millionaire shortbread is a reason for Meghan to be excited about her visit. The fact that Alfred would be delighted is a reason for you to bring him flowers. We assume that the things favoured are states of affairs consisting in particular agents performing actions or having attitudes. The things that do the favouring are not simply objects or properties, but rather particular objects instantiating properties. To avoid a potential ambiguity, we'll say that the fact that *is* or equivalently that *constitutes* the reason is the fact that stands in the first place of this reasons relation. We'll say that the *reasons fact* is the fact that this particular fact favours that agent's action or attitude, i.e. the particular relation itself.

I shall mainly be arguing for the following thesis:

VALUE-BASED THEORY OF REASONS: For some fact to be a reason for you to do something is for that fact to either constitute or partly explain the fact that your doing that thing would realise some state of affairs that is valuable.

Similarly we can say that for some fact to be a reason for you not to do something is for that fact to either constitute or partly explain the fact that your doing that thing would realise some state of affairs that is disvaluable.^{3,4}

The fact that Alfred would be a delighted is a reason to bring him some flowers since your doing so would realise the state of affairs of his being delighted, which is valuable. The fact that Mildred would see the *Adoration of the Magi* is a reason for her to go to the Uffizi because her experiencing that painting is valuable. The fact that the abandoned Great Dane will otherwise be put down is a reason to adopt him because the death of the Great Dane is disvaluable. The fact that distributing free milk in schools would improve equality of opportunity is a reason to distribute free milk in schools because the improvement in equality of opportunity is valuable.

³ For simplicity I will assume that there is no difference between a reason for not-X-ing and a reason against X-ing. This is contentious; see Greenspan 2005 & Gert 2007 for some of the relevant complications. I do not assume that a reason to Y, in a situation in which one cannot X and Y, is *ipso facto* a reason not to X.

⁴ This theory naturally extends to reasons for desires and emotions, though I will not so extend the account here. This value-based theory of reasons does not extend quite so straightforwardly to reasons for belief. I will not deal with this issue here. See Asbjørn Steglich-Petersen 2011 for a recent defence of a value-based account of reasons for belief. Selim Berker 2013 presents some objections to this approach. Some of Berker's arguments turn on claims about practical reasons which are problematic in the light of some of the arguments in this paper, but a full discussion will have to wait for another occasion.

There are various ways in which an action would *realise* some state of affairs that is valuable. Most simply some action would realise such a state of affairs by *causing* it, as in the case of giving Alfred flowers. An action also realises value by *exemplifying* it, as when Mildred's giving Alfred flowers expresses her love for him (on the assumption that expressions of love are valuable). However we need to add that an action also realises value by *probabilifying* it. This is because there are cases in which no value will in fact be caused by some action, despite this outcome being very unlikely. Suppose that for one dollar someone offers you 99 of 100 tickets chosen at random in a fair lottery with a big prize. In fact the 100th ticket will win. Still, intuitively you have a perfectly good reason to pay your dollar for the tickets.⁵ I will use the term 'realises' expansively, to include these three relations to states of affairs.⁶

We are interested in the fact that 'constitutes or partly explains' the fact that your doing the thing would realise some state of affairs that is valuable. By 'constitutes' here I just mean to denote the identity relation. In basic cases, the reason just is the fact that your doing some particular thing would realise some particular state of affairs that is valuable. In less basic cases, the reason will be some fact that, together with other facts, explains some fact like this.⁷ The fact

⁵ Notice the differences between this case and Bernard Williams's drinking petrol case.

⁶ I maintain that some action 'realises' a state of affairs if it probabilifies it, even if the state of affairs does not in fact obtain. This is linguistically infelicitous, but I can find no other English word to play this role. For historical reasons that will come up a little later (see footnote 23) I want to avoid the word 'promote.'

⁷ I have here a metaphysical conception of explanation in mind. For recent work on this notion see Gideon Rosen 201; Kit Fine 1994, 1995 and *forthcoming*; and Paul Audi's 2012. Michael J. Clark and David Liggins have a nice survey of recent work in *Analysis*, 2012. Following most of

that I-93 is backed up for ten miles is a reason to take some other route, since that fact, together with various others, explains the fact that if you take I-93 you'll get very annoyed by the delay. Your being annoyed is disvaluable, we can suppose. The fact that Teresa's tyre is flat is a reason to give her a lift to work, because that fact, together with various others, explains the fact that she will be unable to get to work unless someone gives her a lift. We can suppose that her not getting to work is disvaluable.

We've made a few broad claims about value, so let me make a few remarks about the notion of value that is in play. Our basic building block is the monadic property of *being valuable*.⁸ We will remain neutral about whether this property can be analysed in non-evaluative and non-normative terms – so long as we insist that facts about value are never explained by facts about reasons that are not themselves explained by facts about value.⁹ This property is graded, and so we can generate some sort of (perhaps partial) ordering of states of affairs in terms of amounts of value. We are talking specifically about *neutral* value, i.e. such that your happiness and my happiness may be valuable *in the very same sense*. With some minor adjustments, one could retrofit a value-

these theorists I assume that explanation relates facts. I observe no distinction between facts, states of affairs, or objects instantiating properties or standing in relations.

⁸ There is a debate in the theory of value about whether instead of talk about value and disvalue, we can simply talk about *betterness*, perhaps with some additional assumption about there being some objective 'zero.' (For background, see Broome, J. 1999.) It will be occasionally helpful to talk in terms of betterness, for instance to make sense of the reason to perform the lesser of two evils. My view is compatible with both ways of thinking about this.

⁹ Michael Smith, Derek Parfit, and perhaps Henry Sidgwick have held the view that facts about reasons for action obtain in virtue of facts about value, but that facts about value obtain in virtue of facts about reasons for desires. At least some of these philosophers think that facts about reasons for desires do not obtain in virtue of anything else. On such views reasons are *at the bottom* of the explanatory chain, hence at least some reasons would be prior to values after all. Cf. Henry Sidgwick, 1874; Derek Parfit, 2011; Michael Smith, 2011. I reject two substantive moves in this alternative account, namely that value is explained in terms of reasons for desires, and that reasons for desire are not explained by realisations of value.

relativist account into the general framework outlined here. We allow that states of affairs that have instrumental value or contributory value may play the relevant role in our thesis. These structures have close analogues in the theory of reasons, so this should not be too contentious in the current context.

The value-based theory of reasons is designed to be mostly neutral between different axiologies, that is, between different substantive accounts of which states of affairs are valuable. In particular, the theory is consistent with the thesis that only states of affairs consisting in the performance of actions or the having of attitudes are valuable. In this way the view is friendly to some deontologists.¹⁰ However the theory does involve some general substantive commitments about values, or reasons, or both. The value-based theory of reasons entails the following biconditional, that you have a reason to do something if and only if your doing that thing would realise some valuable state of affairs. The sufficiency claim is very plausible; the necessity claim – that some fact is a reason for you to do something only if your doing that fact constitutes or partly explains the fact that doing that thing would realise some valuable state of affairs – is rather contentious. Two kinds of cases trouble this thesis. Firstly we have ‘deontic reasons.’ Though many reasons seem to be explained by values like fairness, solidarity, well-being, beauty, etc., many others seem to be explained not by values but by rights, obligations, duties, and other deontological-sounding bits and pieces. Secondly we have ‘partial reasons.’ Plausibly everyone has a reason to jump into the pond to save the drowning child, but not everyone has a reason to go to your son’s début ballet

¹⁰ Compare Wedgwood 2009.

performance, or to hope that your partner doesn't get stuck in traffic on the way home.¹¹

An extreme way to respond to these problems is to 'consequentialise.'¹² The idea here is that we take the various rights, obligations and duties which it would behove our theory of reasons to leave room for, and we simply insist that there are values realised by honouring these rights, obligations, and duties. For instance we say it is *ipso facto* valuable to keep a promise. In this way we preserve the deontological intuition that there is always a reason to respect one's commitments and so forth, without upsetting the necessity claim entailed by the value-based theory of reasons. Something about this approach must be right. We should be open to expanding our axiology to include new kinds of values. However we should not do this in an *ad hoc* or arbitrary way. Rather we should be actively working to find values which explain our deontic intuitions. These will often involve notions of trust, reciprocity, fairness, dependence, and so on.

One might worry that this account does not quite preserve the normative significance of deontic reasons and partial reasons, and that it cannot deal with certain *symmetric* cases, in which more of the same value will be realised by the agent's allowing others to perform partial actions. We can mitigate this worry

¹¹ Perhaps also, thirdly, there is a problem specific to the value-based theory of reasons for attitudes. In many cases, it will seem plausible that one has reason to have a certain attitude, and equally plausible that having that attitude will not be *instrumentally* valuable. To accommodate such cases it looks like we need to insist that these attitudes instantiate value; fortunately this seems independently plausible. See chapter two.

¹² On the now-familiar notion of consequentialising, see Bergstrom 1966; Broome 1991; Sen 1982; Vallentyne 1988; Dreier 1993; Louise 2004; Smith 2009; Brown 2011.

with a variation on this approach, which I defend at length in chapter two. The idea here is that familiar values ground reasons, even if only small amounts of value are at stake. This approach adds that the weights of these reasons can be effectively intensified by the sorts of commitments and relationships that the deontologists are concerned with. This approach combines the plausible denial that there are reasons in valueless cases, with the plausible insistence that partial reasons and deontic reasons are distinctively weighty. A nice feature of this alternative approach is that it shows that value-based theories of reasons need not be consequentialist. So long as the weights of reasons are not directly proportional to the amounts of value that explain them, we leave room for the possibility that in some, perhaps many, cases one will have more reason to do something that would realise less value than some alternative.¹³

SECTION TWO: THE BUCK-PASSING THEORY OF VALUE

We can now consider the most prominent challenge to this value-based theory, namely T.M. Scanlon's view, as developed in *What We Owe To Each Other*. Scanlon's popular 'buck-passing' theory of value attempts to analyse the property of *being valuable* in terms of reasons. His view was introduced in opposition to value-based theories of normativity like mine (albeit ones concerned with *ought* rather than with *reasons*).

¹³ It is worth pointing out that buck-passers also has problems explaining partial reasons. We'll talk about this more in section three. See also Zimmerman 2011 and references therein.

Although Scanlon did not draw this distinction himself, it is now well-known¹⁴ that he in fact introduces two buck-passing theses, a negative and positive thesis. According to the negative buck-passing thesis, “being good, or valuable, is not a property that itself provides a reason to respond to a thing in certain ways.” According to the positive buck-passing thesis, “goodness and value [are]... the purely formal, higher-order properties of having some lower-order properties that provide reasons of the relevant kind.”¹⁵

The negative buck-passing thesis does not entail the positive buck-passing thesis. This is clear since it is logically possible that there is no correlation between values and reasons; for instance this was W.D. Ross’s view in *The Right and the Good*. We can add that the negative buck-passing thesis is ambiguous. The idiom ‘provides a reason’ is ambiguous as between ‘constitutes the reason’ and ‘is part of the full explanation for the existence of the reason.’ Scanlon seems to have the first in mind,¹⁶ i.e. that the fact that some action or state of affairs is valuable does not constitute the relevant reason.

It is important for current purposes to notice that I can also accept this thesis that facts about states of affairs being valuable do not constitute reasons. In my view, the reason is the state of affairs that is valuable, not the state of affairs consisting in that very state of affairs being valuable. It is the fact that Alfred

¹⁴ On the distinction between the positive and negative theses, see Schroeder 2011; Väyrynen, 2006, p. 296; Crisp 2008, p.263; Heuer (ms), p.1.; Dancy, 2000 & 2003.

¹⁵ 1998, p.97.

¹⁶ This is implied by the use of the same word ‘provides’ in Scanlon’s characterisation of the positive buck-passing thesis, and by his earlier characterisation of the view to which he is opposed: “[the] property of being valuable...*gives us* reason to behave or react in certain ways...” (ibid; my italics.). Ulrike Heuer (ms), for instance, explicitly characterises Scanlon as accepting the first of these.

will be delighted, that Mildred will be able to see the *Adoration*, and so forth, that constitutes the reason. The additional fact that Alfred's delight is valuable *explains why* this fact constitutes a reason.¹⁷ Strictly speaking, we say that the fact that your giving Alfred flowers would delight him is a reason in virtue of the fact that his being so delighted is valuable.¹⁸ This all strikes me as rather natural and intuitive. It helps the value-based theory of reasons to avoid the sort of alienation objections that bedevil consequentialism.¹⁹ But more pertinent to the current context, this structure takes some of the force out of Scanlon's main argument against value-based theories. Since we can both accept this thesis, i.e. the thesis that the fact that some action realises value is not itself a reason to do it, let us henceforth put this 'negative' buck-passing thesis to the side.

The positive buck-passing thesis provides an account of value in terms of reasons. Roughly, the idea is this: For something to be valuable is for it to have properties such that the fact that the thing instantiates those properties constitutes a reason for agents to respond positively towards it. Here we have

¹⁷ Compare Schroeder 2007, p.14: "There can be conditions that have to be satisfied in order for something to be a carrot, which don't count as part of the carrot." For more on this point, see also Brad Hooker & Philip Stratton-Lake 2006, p.161 and Pekka Väyrynen 2006, p.321.

All that being said, it doesn't seem *obvious*, as Hooker and Stratton-Lake suggest it is, that facts about value cannot directly provide reasons. It is an additional advantage of my own view over the buck-passing view that it is consistent with either the truth or falsity of this thesis.

¹⁸ Compare Raz 1999, p. 23. Dancy misses this possibility in his five-option taxonomy in 2000, p.164 & 165. (It is most similar to his option 4, if you swap the 'v' and the 'r'.)

¹⁹ Williams, B. 1973 and Railton, P. 1984 are the central texts discussing the alienation worry. My approach here owes a debt to Mark Schroeder's response to the 'self-regarding character' objection, in 2007, p. 27. As we shall see, the value-based theory of reasons is at two removes from this alienation worry. Firstly it gives all the various realisations of value their own normative status, precisely as reasons, rather than simply giving normative status to a single maximising option. Secondly as we saw in the main text, the distinction between the facts that constitute the reasons, and the facts in virtue of which they are reasons, and the further facts that these facts that constitute the reasons are reasons, or stand in the reason relation, allow us to say that the very token realisations of value constitute reasons. So it is the fact that this girl will be hurt, or that you promised Joe, or what have you, that is the reason to do this or that.

facts about value being explained in terms of facts about reasons. Given the asymmetry of explanation, this positive buck-passing thesis is clearly inconsistent with the value-based theory of reasons. Henceforth I'll refer to 'positive buck-passing' just as 'buck-passing.'

Buck-passers do not have to be reasons fundamentalists. They are free to give an analysis of reasons in other terms. They must simply insist that facts about reasons are not fundamentally explained by facts about values.²⁰ For simplicity, however, we'll assume that buck-passers are reasons fundamentalists (i.e. they take it that the relation of being a reason cannot be analysed in terms of other notions) and the defenders of the value-based theory of reasons are value fundamentalists (i.e. the same for the property of being valuable).

SCANLON'S FIRST ARGUMENT: FROM EXPLANATION

Scanlon has two arguments for the buck-passing theory of value. Here is the first:

[W]hen I consider particular cases it seems that these reasons are provided by the natural properties that make a thing good or valuable.

So, for example, the fact that a resort is pleasant is a reason to visit it or

²⁰ By 'fundamentally explained' I mean to allow that some facts about reasons might be explained by facts about value, but that whenever that is the case, there is always some further fact about reasons which explains this fact about value. (Perhaps this in turn is explained by some further fact about value – but facts about reasons will always be at the bottom of such explanatory chains.) The structurally analogous point holds for value-based theories – see fn 8 above.)

to recommend it to a friend, and the fact that a discovery casts light on the causes of cancer is a reason to applaud it and to support further research of that kind. These natural properties provide a complete explanation of the reasons we have for reacting in these ways to things that are good or valuable. It is not clear what further work could be done by special reason-providing properties of goodness and value, and even less clear how these properties could provide reasons.²¹

This can easily be framed as a challenge to the value-based theory of reasons as we are understanding it. We all agree that certain facts – whether natural facts or thick facts – constitute reasons. I claim that some further fact explains why those facts and not others constitute reasons, namely because they stand in the relevant relation to a state of affairs that is valuable. Scanlon denies that this further step is necessary. We can reformulate this into a challenge: what extra work is left undone by appeal to the facts that constitute the reasons? Why do we need this appeal to the property of being valuable?

I will attempt to meet this challenge in the rest of this paper. For now, however, notice that Scanlon faces a structurally similar challenge. We are comparing two different accounts of reasons and values. One takes the property of being valuable as fundamental (for the sake of argument) and attempts to explain facts about reasons in terms of it, together with other bits and pieces. The other takes the reason relation as fundamental (for the sake of argument) and attempts to explain facts about value in terms of it, together

²¹ 1998, p.97.

with other bits and pieces. So I can pose my own explanatory challenge: what extra work is left undone by the fact that one's action or attitude would realise some state of affairs that is valuable? Each of these accounts must be assessed on its merits both as a theory of value and as a theory of (practical) reasons. I'll be arguing that the value-based theory is explanatorily superior in a number of respects.

SCANLON'S SECOND ARGUMENT: FROM PLURALISM

Scanlon's second argument runs as follows: "...many different things can be said to be good or to be valuable, and the grounds for these judgements vary widely. There does not seem to be a single, reason-providing property that is common to all these cases" (1998, p.97,98). Scanlon takes this to support his buck-passing view indirectly, by constituting a reason to reject the main alternative to his account.²² Shortly afterwards he adds:

Once one recognizes the variety of things that can be valuable and the variety of responses that their value calls for, it becomes highly implausible that there could be a systematic "theory of value." Understanding the value of something is not just a matter of knowing how valuable it is, but rather a matter of knowing how to value it – knowing what kinds of actions and attitudes are called for. It is an

²² He does not consider the Rossian alternative that neither values nor reasons explain each other, nor the other positions that Dancy 2000 suggests.

advantage of the present account that it calls attention to this aspect of our ideas of value, one that is easily concealed by the assumption that the primary question about the value of something is how great that value is.

These passages suggest three different points. We can take this as an argument against (a) the substantive thesis that there is only one, or perhaps a very small number, of different types of states of affairs that are valuable. But as Philip Stratton-Lake and Brad Hooker (2006) point out, the substantive debate in axiology between monists, pluralists, and particularists is strictly orthogonal to the merits of the buck-passing account. Hence it is irrelevant to our dispute. Scanlon certainly means to cast some doubt (b) on the hypothesis that there is some single property – *being valuable* – instantiated by all and only the things that provide reasons. This is where our earlier distinction is important. The states of affairs that provide reasons are various; but they may all instantiate the *further* property of being valuable. It is a separate question whether this further property explains or is explained by the fact that the states of affairs that instantiate it partly constitute reasons.

Finally, what I think is most prominent in Scanlon's mind in his argument from pluralism is (c) monism about the sort of *responses* to valuable states of affairs that are appropriate. He is specifically concerned to oppose the idea that the only appropriate response to value is to try to maximise it. Nearly twenty years after *What We Owe To Each Other*, I think we can say that this lesson has been learned. Some theorists still try to preserve a substantive distinction between

consequentialism and deontology by means of a distinction between ways of responding to value.²³ I think this is a mistake. We have reasons to respond to things, valuable and otherwise, in many different ways. In order to explain why we have reason to react differently to Ming vases and atom bombs, Scanlon will appeal to facts about Ming vases and atom bombs. But these resources are also available to the value-based theorist.²⁴ In fact, as I'll now suggest, these facts about the plurality of reasons for responding to things actually present important difficulties for the buck-passing theory.

SECTION THREE: ARGUMENTS FOR THE VALUE-BASED THEORY OF REASONS

ARGUMENT ONE: STRUCTURAL ADVANTAGES

We stated the buck-passing thesis above as follows: for something to be valuable is for it to have properties such that the fact that the thing instantiates those properties constitutes a reason for agents to respond positively towards it. It is well-known that at least two restrictions need to be added to this analysis in order for it to be *extensionally* plausible. I'll suggest that there are in fact four restrictions that need to be added, and that the need for these undermines Scanlon's putative *explanatory* advantage.

²³ Consequentialists such as Pettit moved from 'maximisation' to 'promotion' under pressure from examples like friendship, the value of which does not seem apt for maximisation. Pettit distinguishes promotion from honouring, which latter he claims is the distinctive deontological attitude towards value. My approach is more ecumenical. Insofar as honouring value, or honouring anything – false idols, if you like – realises value, there is a reason to do so.

²⁴ It is worth mentioning, *ad hominem*, that Scanlon often precisely appeals to evaluative language in his explanations of why certain facts constitute reasons. For example see 1998, p. 38, 49, 65-69, 150, 200, 513; and 2003, p. 96. For worries about this, see Pekka Väyrynen 2006.

The best known problem for buck-passing has become known as the ‘wrong kind of reasons problem.’²⁵ The fact that the gunman will shoot you unless you have positive attitudes towards panpipe music gives you a reason to do so but it doesn’t make this music valuable. To fix this, we need to restrict the analysans to some particular account of the *right kind* of reasons.

A second problem, noted by Roger Crisp,²⁶ is that of distinguishing the *specific attitudes* reasons for which plausibly explain the fact that something is valuable. With general attitudes, like ‘choose’ or ‘prefer’ we get false positives. More specific pro-attitudes like ‘admire’ or ‘respect’ perhaps avoid this problem, but perhaps only because we understand these attitudes by way of facts about admirable or respect-worthy objects being valuable in some significant respect. The buck-passer must identify something in common among all the appropriate attitudes that ground value. This puts pressure on the buck-passer’s argument from pluralism. But more worryingly here, the buck-passer needs to do this without furtive appeal to the fact that the things in question are valuable.

An interesting third problem, first raised by Brand Blanshard as an objection to A.C. Ewing’s buck-passing account,²⁷ is that different people have reasons to react to the same things differently. For instance I have reasons to be partial towards my family, or towards books I grew up with, and to be averse to poems or songs which remind me of unpleasant past events. I lack reasons to

²⁵ For a helpful bibliography, see Jussi Suikkanen 2009.

²⁶ In his 2005 and 2008.

²⁷ For references and discussion, see Zimmerman, 2011.

be excited by developments in genetic biotechnology, since I lack the sophistication necessary to understand them. The buck-passer may reply that I have reasons to be excited by these developments, but these reasons are weak. But this suggests a further problem for the analysis. It is a plausible desideratum for a theory of value that it explain not simply the fact that something is valuable but also how valuable it is. The most natural way for a buck-passer to explain these facts would be to appeal to facts about the *weights* of the reasons in question. But the variability of different people's reasons presents a significant obstacle to this.

The fourth problem is related. *Prima facie* some state of affairs might be valuable or disvaluable, even if there is no-one around to have reasons to react to it one way or another. Krister Bykvist presents the example of "there being happy egrets but no past, present, or future agents."²⁸ The idea is that the state of affairs consisting in these egrets being happy has some positive amount of value. For instance, it is much better than the state of affairs consisting in there being egrets suffering extreme amounts of pain, but no past, present, or future agents. But since there are no agents, no-one has any reasons to respond to this situation.

On behalf of the buck-passer, Jonathan Way replies as follows:²⁹

²⁸ 2009, p.5.

²⁹ 2013, p.8.

“...the fact that there can be no agents in this state of affairs is not a good reason to think that there can be no agents suitably related to this state of affairs. (Nor does Bykvist claim otherwise.) The notion of a ‘suitably related’ agent which is relevant here is psychological and perhaps also epistemic. In this sense, for example, we are not suitably related to, and so lack reasons to have attitudes towards, outcomes which we lack the concepts to conceive of. But there is nothing to stop agents being in this sense suitably related to states of affairs that they are not a part of. For example, in Bykvist’s case *we* have reasons to have attitudes towards the state of affairs of there being happy egrets but no past, present, or future agents – for instance, reasons to be glad that this state of affairs does not obtain.”

This response is fine so far as it goes, but it involves a further restriction to agents *suitably related* to a state of affairs. So at the end of all this, the buck-passing thesis is as follows:

BUCK-PASSING THEORY OF VALUE: For something to be valuable is for it to have properties which provide reasons of kind R with weight W for agents meeting condition C to have attitudes of kind A towards it.

These various restrictions have generally been presented as extensional problems for buck-passing, which of course they are. But they also undermine the theory’s explanatory credentials. Suppose for now that the buck-passer provides us with a satisfactory account of the right/wrong reasons distinction

perhaps along with an account of appropriate weight, an account of what it is to be 'suitably related' to a state of affairs, and an account of which pro-attitudes are doing the explanatory work in different cases. Consider two states of affairs: (1) happy egrets but no past, present, or future agents; (2) sad egrets but no past, present, or future agents. Our value-based theorist will say it is a simple axiological fact that 1 is better than 2. The explanation for this simple axiological fact just has to do with the happiness and sadness of the egrets. The buck-passer must say that the fact that 1 is better than 2 obtains in virtue of some very complicated pair of facts about suitable agents having the right kinds of reasons to respond to these states of affairs with the right kind of attitudes to the right degree.

This pair of explanations is the complement of the explanations contrasted by Scanlon in his argument from explanatory adequacy. The simple fact for Scanlon was the fact that the pleasantness of the resort provided a reason to visit. The more complex fact was that the fact that going to the resort would be pleasant constituted a reason, in virtue of the fact that experiencing pleasant resorts is valuable.

But the explanations on each side are not equally good. The value-based theory is straightforwardly constructed from the monadic property of being valuable together with the *realises* relation; this perspicuity is a *prima facie* advantage over the promissory nature of the buck-passing analysis. Moreover, importantly, it is very plausible that facts about value play the role of sorting those facts that do from those facts that do not constitute reasons, at least in many cases.

Plausibly, those cases in which we feel that the appeal to value is otiose are simply those in which it is *epistemically* redundant; it is plausibly still part of the full explanation in accordance with the value-based theory of reasons. We find some further evidence for this in the fact that Scanlon himself often reaches for evaluative language when explaining why some fact does or does not constitute a reason.³⁰ By contrast, absent further theoretical motivation, it is implausible that the fact that 1 is better than 2 obtains in virtue of the long complicated fact about suitable agents, right reasons, and so forth. But there is no such further theoretical motivation: the arguments from pluralism and explanation provide no support for this claim. Hence it is the value-based theory, and not the buck-passing theory, which has the overall explanatory advantage.

There is a further worry here, namely that there is a simple debunking hypothesis which explains why buck-passing encounters all these various difficulties. The hypothesis is this: buck-passing is really a theory of appropriate valuing, not a theory of value. This hypothesis generates further counterexamples to the buck-passing account: for instance it can be appropriate to value something if someone we respect values the thing, or if doing so would thereby make it valuable though it isn't yet (as with a social institution for instance), or if doing so would inculcate valuable dispositions. Perhaps these are wrong kinds of reasons of a sort, but they are clearly different from the traditional 'value it or I'll shoot' cases. This hypothesis would also explain why buck-passers have to work hard to avoid getting the wrong result in cases involving abstruse things of value, or things of value which we have

³⁰ See footnote 24 for a lengthy list of references.

idiosyncratic reasons to ignore. This all casts doubt on the prospect of some eventual refinement of the buck-passing analysis even achieving *extensional* adequacy, without furtive explanatory reliance on facts about the values of states of affairs.

ARGUMENT TWO: COUNTING REASONS

The fact that Teresa's tyre is flat and the fact that she can't get to work on her own are both reasons for me to give her a ride, but they are not really distinct reasons. The fact that she offered me \$100 to give her a ride is a distinct reason. Going to the dinner tonight would give me an opportunity to meet Bonecrusher Mike. That's a reason to go. I need to make a deal with Bonecrusher Mike. That's a reason to go. But these are not really distinct reasons; compare the fact that they will be serving swordfish. "The fact that there will be dancing at the party is a reason for Ronnie to go there. The fact that Ronnie likes to dance is a reason for him to go to the party."³¹ But these are not really distinct reasons; compare the fact that Jessica will be there.

There is a puzzle here. When are reasons *distinct*, and when do they *overlap*? The point here is not exactly about the individuation of reasons; it is about figuring out when facts contribute separate weight. It is important to give an account of when non-identical reasons overlap in this way for various reasons. Firstly this is a feature of our ordinary reasons-talk, which any complete theory

³¹ Schroeder, M. 2011, p. 3.

of these matters will have to explain. Secondly the notion of overlap under consideration will be helpful for arbitrating no-relevant-difference arguments. Such arguments depend on claims of the form ‘the reason not to do this is the same as (overlaps with) the reason not to do that, and you agree that there is a reason to do that.’ Since this is perhaps the most powerful argument form in the ordinary person’s normative canon it would be valuable to have some criterion for assessing such claims. Thirdly, and perhaps most importantly, an account of this phenomenon of overlapping reasons will provide us with resources with which to avoid double-counting; in other words it is important for the theory of weighing reasons, and hence to one’s theory of the all-things-considered ought.

The value-based theory of reasons has a straightforward solution to this problem: for two reasons to overlap is for them to be explained by the same realisation of some particular value. This criterion explains the overlapping reasons in the Bonecrusher Mike, the Teresa’s car, and the Ronnie’s dancing cases. It is the value of making the deal with Mike, or of getting Teresa to work on time, or of Ronnie shaking his tail-feathers, that explains why the reasons in question overlap.

Scanlon’s explanatory resources are exhausted by the facts that constitute reasons and the reasons facts themselves. As we have seen he can explain different reasons by appealing to basic reasons. This suggests an alternative account: for two reasons to overlap is for them to be explained by the same reason, or for one of them to be a basic reason and for it to explain the other.

However this is interestingly less specific than the value-based proposal. The value-based account appeals to particular realisations of value; the reasons-based account appeals to the reasons at the bottom of an explanatory chain. This means that the reasons-based account would end up getting the wrong result in certain cases. Suppose I promise Joe that I'll take Barbara out for lunch, and later I promise Frank that I'll take Barbara out for lunch. I have two non-overlapping reasons here, but both are explained by the same basic reason, namely, the fact that taking Barbara out for lunch is a way of keeping my promise. Or suppose three good long-lost friends of yours will be at the bar tonight. That's three distinct reasons to go to the bar. But plausibly they are reasons to do the same thing explained by the same basic reason or reasons, namely the fact that by going to the bar I will be able to spend time with an old friend. Or consider a third example. The fact that millions of children are dying of preventable diseases in Sub-Saharan Africa is a reason for me to donate a proportion of my income to charity. The fact that there is a drought in Southern India is another reason for me to donate a proportion of my income to charity. These reasons are plausibly explained by the same basic reason: the fact that by donating to charity I will be doing what I can to save innocent lives.

These examples suggest that overlap does not coincide with explanatory depth. These examples are specific in one particular way. In each case overlap is explained by some *particular object* instantiating a general property: this friend or that friend, these children, Barbara. To accommodate this the reasons fundamentalist can either abandon the idea that more general reasons explain

more particular reasons, or look for another criterion for overlap. I can see no non-ad hoc reason for either.

Another potential problem arises when we think about the fact that reasons *for different people* can overlap. Consider some further cases. Five of us have keys to the apartment. Jacob is locked out. The remaining four of us have a reason to go back to the apartment as soon as possible. But we are about to sing as a barbershop quartet. Perhaps each of us wants to sing for different reasons – one to please her partner, one to revitalise the genre, and so on. Our four reasons to go back to the apartment do not add up separately, to counterbalance the reasons to sing. They overlap. Intuitively we have one reason to go home, along with whatever reasons we may have to sing. Or consider: Teresa needs a ride to work, and you or I could help her. But we both want to go for a jog. Intuitively in this case we have one reason to help Teresa, and two separate reasons to jog. Our reasons to help Teresa overlap.

Again the value-based theory provides a natural explanation of this. One valuable state of affairs will be realised by getting Teresa to work, no matter who does it. That explains why our two reasons do not add up. By contrast two distinct valuable states of affairs will be realised by our jogging, namely your jogging and my jogging. These should contribute their own separate weight.

Here the buck-passing view can appeal to the most fundamental reason in the explanatory chain only if such ‘agent-neutral’ reasons indeed explain reasons

for particular individuals. This is a claim some buck-passers have wanted to deny, perhaps because of the worry about appealing to principles, noted above. Scanlon, for instance, is somewhat unclear about this, claiming that universalisation is a matter of a condition on reasons *judgements* rather than reasons themselves.³² My point here is that the same sorts of considerations that push a buck-passer to explain reasons for individual agents in terms of reasons for all agents should push her to explain reasons for particular states of affairs in terms of reasons for general states of affairs. So the buck-passer cannot comfortably explain these cases *and* the previous set of cases.

In the face of these two sets of cases, a reasons fundamentalist can reject the 'bottom of the explanatory chain' criterion for counting. She can simply stipulate that reasons of the right sort are to be counted (that is, generally-described agents realising states of affairs consisting of particular objects instantiating general properties). However, although this gets the cases right, it lacks internal motivation within the theory. So it looks as though buck-passers either get the wrong result, or their view is somewhat ad hoc.

By contrast, the value-based theory offers a straightforward explanation of this result. Valuable states of affairs consist of particular objects instantiating fairly general properties. These will be the properties featuring in axiological principles, assuming there are such things – and still not taking a stand on the question of whether such principles explain or are explained by their instances. It is the fact that Mildred is happy, or that Frank expresses solidarity with

³² 1998, p.74

Charles, or that the mango tastes delicious, that is valuable. These particular valuable states of affairs explain facts about overlap.

ARGUMENT THREE: WEIGHING REASONS

Reasons can be compared in terms of their relative strength or weight. Your reason to meet Gerald for lunch is less weighty than your reason to attend your daughter's graduation, and that less weighty than your reason to defuse the bomb set to blow up Tokyo. Alfred's reason not to open his umbrella in the lightning storm is weightier than his reason to do so. Perhaps Paul Gauguin's reasons to stay with his family were weightier than his reasons to travel to French Polynesia. Mark Schroeder (2007) taught us to not assume that our theory of the existence of reasons will explain the weight of reasons. But one's theory of the existence of reasons will importantly *constrain* one's theory of weight.

According to the value-based theory of reasons, non-basic reasons are explained by basic reasons. Moreover all of the reasons explained by some given basic reason to X will contribute exactly the same amount of weight in favour of X, namely precisely the amount of weight contributed by the basic reason that explains it. Distinct reasons explained by the same basic reason (e.g. the fact that Teresa's tyre is flat and the fact that she can't get to work on her own) do not contribute *separate* weight to X-ing.

We can now add that there is a correlation between weights of reasons and *amounts* of value likely to be realised. Suppose that Jack promised to bring milk home, or to meet Jill for dinner, or to defend you in court, or to “faithfully execute the office of President of the United States.” We assume that these promises all generate reasons. There are clearly different disvalues likely to be realized by failing to satisfy these promises. The weights of the reasons generated by these promises are correspondingly different. Other things equal, and assuming he could not do both, the reason to defend you in court would outweigh the reason to bring milk home. Now suppose Frank can help Barbara across the road, or carry her shopping home, or do her taxes, or donate her a kidney. These things are increasingly valuable. Consequently the weights of the reasons to do these things increase as we go along the list. (Of course there may also be increasingly weighty reasons not to do them. The pain and inconvenience to Frank of donating a kidney, for instance, must be balanced against the value to Barbara.)

We find support for this thought in some wonderful passages from one of the early papers specifically concerned with the weight of reasons³³:

“...consider truth-telling. Lying in scientific contexts (for example, by falsifying data) is obviously a serious matter since the truths at issue are important ones and it does not take much fraud seriously to jeopardize the scientific enterprise. Lying to strangers that one is unlikely to meet again about trivial matters is far less serious since the truths at issue are

³³ Michael Philips, ‘Weighing Moral Reasons,’ 1987

of little importance to him and since trust between strangers is not seriously endangered. If anything is wrong with such lies, it is that they might promote a habit that spills over into other contexts. In general, the point of the consideration in favour of truth-telling is to facilitate the exchange of useful information and to facilitate relations of trust between people. The weight of that consideration increases with the importance of the information exchanges and with the degree to which important forms of trust are endangered by its violation.

Corresponding considerations hold in relation to promise-keeping, stealing, and so on. One purpose of the rule against stealing is to make us secure in our property so that we can put it to the uses we have planned. For this reason, other thing equal, the wrongfulness of stealing is greater when it is reasonable to believe that the victim will be more inconvenienced by the theft. Thus, for example, other things equal, it is worse to steal \$100 from a poor person than to steal \$100 from a rich person. Similarly, one good of promising is that it enables us to count on the others to do as they say, and hence to cooperate with them in various activities and institutions. For that reason, other things equal, the most important the activity or institution made possible by promising, the more seriously wrong it is to break a promise. Thus, it is more seriously wrong to violate marriage vows or a professional oath than to break a promise to meet a friend for lunch.

Weight is a function of the amount of value *likely* to be realised. The degree of likelihood constitutes a further factor in the weight-determining function. There may be further factors. For instance there might be a factor for the relation between the agent and the value in question: whether it is value to herself, to a friend, or to a total stranger; whether she has made some *commitment* to realising the value in question; whether the realisation of values like this is one of her *projects*. In this way we make theoretical room for *partiality* to affect the weight of reasons. Hence while you and I may have the same reasons to save the rhinos, or to rally in Kelvingrove Park, or to watch your daughter's *début* ballet performance, your reasons may be stronger.

This is all rather natural and attractive. It has been clear since long before David Hume wrote 'Of the Standard of Taste' that while we may have difficulty making comparative value judgements on the margins, we have no difficulty at all in many cases ("...whoever would assert an equality of genius and elegance between Ogilby and Milton..."). The value-based theory of reasons takes as fundamental a notion of value which is *ordered*, at least in the sense in which any two states of affairs are either better than, worse than, the same as, or perhaps on a par with each other.³⁴ Lots of work has been done in axiology, in economics, and in decision theory to straighten out our intuitions about these matters.

But of course we also have intuitions about the relative weights of reasons, which may in time be systematised just as usefully. So while the idea of a

³⁴ Cf. Chang, R. 1997.

fundamental amount of value may strike us now as more tractable than the idea of a fundamental weight of reasons, this might be misleading. Making judgements about the relative clarity of intuitions about putatively fundamental properties is a murky business.

The most natural option for a reasons fundamentalist is to build *weightiness* into his fundamental *reasons* relation. Thus instead of a four-place relation between fact, agent, situation, and action we have a five-place relation, with an extra place for weight.³⁵ There is plenty of precedent for notions which might be fundamental but which come in degrees: we have the *strength* of a desire, *degree* of credence, and *amount* of value.

This account of the weight of reasons may have to do double-duty, also explaining the amounts of values which are grounded by the reasons in question. Strictly speaking, the buck-passer is giving an account of what it is for something to be valuable, rather than what it is for something to be valuable *to some degree*. But it is natural to think that if value is explained in terms of reasons, then amount of value should also be so explained. Again the most natural choice will be to explain amount of value of some state of affairs in terms of the weights of the reasons to act positively towards it. This puts additional pressure on the intuitiveness of Scanlon's approach. It was already a

³⁵ Cf. John Skorupski 2010 develops such an account. This is not Scanlon's preferred theoretical option. In drafts of his forthcoming *Locke Lectures* he introduces a new fundamental relation, that of a sufficient reason, alongside the usual (pro tanto) reason relation, and uses this new relation to explain the relative weights of reasons. Since this work is not yet in print, I will reserve judgement about it here.

strain to accept that the disvalue of the sadness of the egrets is to be explained in terms of reasons to for hypothetical agents to feel sad for them, or whatever the appropriate attitude is supposed to be here. It is even harder to believe that the amount of disvalue of the egret's sadness is in turn explained by the weight of these reasons, whatever they are, where the weight of reasons is a fundamental matter.

If we take seriously the theory of weighing sketched above, this leads to one further structural problem for Scanlon's view. The suggestion above is that the weight of a reason is a function of more than just the value of the state of affairs that would be realised. It is also a function of the likelihood that the action will realise that state of affairs, perhaps along with facts about the agent's level of involvement in the outcome, and facts about the agent's relationships, commitments, projects, and so on. Some of these quantities will be available to the buck-passer. Facts about likelihood are probably non-normative and non-evaluative. So she can allow that weights of reasons might covary with likelihood – albeit this redirects our gaze to whatever it is that is being made more likely. Some of the facts about the agent's relationships will be non-evaluative. But plenty of the facts about the agent's relationships, commitments, and projects are evaluative. It is precisely the value of the agent's relationships that determines the extent to which they affect the strength of her reasons. On a traditional consequentialist view, these values would simply be summed along with the rest. But on the view sketched above, these values play a different role; they allow for the possibility that we might have weightier reasons to do something that brings about less value, for

instance you might have weightier reasons to save your mother rather than some more beneficent stranger. But now we have one quantity – the weight of reasons – varying with two quantities – the value of the outcome and the value of the relation to the outcome. It is this feature that presents problems for the buck-passer. This will lead to blind spots, just as if we were trying to work out facts about heights from facts about height differentials. The buck-passer will be unable to distinguish two equally weighty reasons, in one of which less partial value is realised, and in the other of which more impartial value is realised (for instance 4 units of pain to your mother versus 7 units of pain to a stranger). It will not be able to distinguish two equally weighty reasons, in one of which less disvalue is realised by more agential involvement, in the other of which more disvalue is realised by with less agential involvement (for instance walking past a girl drowning in a pond versus not contributing to charity). Since the value-based theory explains weight as a function of amount of value realised, along with these other factors, it can easily account for these differences.

We can also make the converse point against the buck-passing thesis. Imagine someone who really likes mud. Loves the feel of it. The fact that there is a huge mudslide on the side of the road is a reason for her to want to stop the car and get into her swimsuit. But it doesn't make the mud valuable. Perhaps it make the mud valuable-relative-to-her, or something like that. But the buck-passer aims to analyse *neutral* value not relative value. We get the same result with partial reasons: think about any mother's love for her evil son. In order to explain neutral value, then, perhaps we have to restrict the account to reasons

that *anyone* would have to respond positively? But this will be too restrictive. I don't have reasons to get excited about the discovery of some new subspecies of beetle (I am terrified of beetles), or about art that I don't like for idiosyncratic reasons: perhaps my first girlfriend dumped me during a rendition of the putatively beautiful song. Perhaps the buck-passer will respond that I do have reasons in these situations, but they are rather lightweight. But this is to jump out of the frying pan and into the fire, for the buck-passer must not just explain value but also amounts of value. The most natural explanans for doing so are the weights of reasons for responding positively: the weightier the reasons, the greater the value. But the weights of reasons for individuals to respond seem plausibly to be affected by *additional* facts, such as their history with or preferences regarding this type of thing. This variability in our reasons to respond to things gives us a defeasible reason to favour the value-based theory over the buck-passing theory.

To summarise. It is very plausible that weight is a function of amount of value and other things, e.g. likelihood, partiality, quality of evidence. It is not plausible that amount of value is a function of weight of reasons together with other things. Hence we should think that facts about amounts of value (together with further facts) explain facts about weights, and not the other way around.

CONCLUSION

Among other things, *What We Owe Each Other* aimed to politely undermine a long and illustrious history of value-first normative theory. A glance at the last fifteen years' worth of the relevant literature reveals that it has been largely successful. It is time to resist this a little.

The value-based theory has the same resources as the buck-passing theory to explain the plurality of types of facts that constitute reasons for us, and the variety of types of responses that may be appropriate. Moreover the value-based theory has several explanatory advantages. It is more plausible that facts about value obtain just in virtue of non-evaluative and non-normative facts, than that they obtain in virtue of the right hand side of the final statement of the buck-passing analysis. This property of being valuable, together with the realises relation, explains facts about reasons simply and powerfully. By contrast, the fact that this and not that constitutes a reason for something seems to many to call out for an explanation, which it would find in the fact that this and not partly explains the fact that doing the thing would be valuable, or at least better than the alternative. The explanation of states of affairs being valuable, in terms of the right hand side of the final statement of the buck-passing analysis, looks very implausible in the case of the lonely egrets, and strikes me as no more plausible anywhere else. Having facts about value at the bottom of our explanations of reasons enables us to count reasons, a matter of great importance when we come to figuring out the overall balance of reasons. It also enables us to explain facts about the weights of reasons. Such facts

appear to be explained in terms of facts about value together with various other facts, such as likelihood. By contrast facts about amounts of value do not seem to vary with weight of reasons and other facts. This backs up the thought that facts about value are more directly explained by features of the states of affairs that are valuable, and that facts about reasons are explained in terms of these together with facts about relations between agents and these states of affairs.

CHAPTER TWO: A NEUTRAL VALUE-BASED THEORY OF PARTIAL REASONS

ABSTRACT

There are several powerful motivations for value-based theories of normativity, of which the various forms of consequentialism are the most famous instances. However value-based theories of normativity, in the form of these most famous instances, encounter serious difficulties accounting for a number of pervasive and important ethical phenomena. These phenomena all involve partiality of one form or another; specifically they concern one's personal relationships, personal projects, and ethical and voluntary commitments. I argue that these difficulties result from an overly simplistic conception of the resources available to value-based theories of normativity, and specifically from an overly simplistic conception of the relation between various kinds of facts about value and facts about the weights of reasons. I present and defend an alternative ethical principle which not only accommodates our intuitions about partiality, but fully vindicates these intuitions. According to this principle, the weight of a reason for you to do something is an increasing function of the amount of value that explains the reason in question and the value of your having a pro-attitude towards this value. I argue that this principle preserves what is distinctively

attractive about prominent forms of consequentialism (act, global, and indirect consequentialism), while avoiding their distinctive shortcomings. The principle also accommodates attractive features of deontology and virtue ethics.

SECTION ONE: THE PUZZLE

You are walking past a pond, and you see a little child caught in the weeds, drowning. You have a reason to help her, because otherwise she'll drown. There is something objectively bad about the drowning of this little child in the pond. Intuitively, it is bad in and of itself. Its badness does not depend on anything to do with you or me. It has everything to do with the child, her suffering, her loss of life-prospects, and so on. These facts about *her situation* give you or me reasons to help her.

Similarly when you see that Mildred is struggling with her bag up some stairs, or realise that Alfred needs cheering up. You have a reason to help Mildred with her bag up the stairs, and you have a reason to give Alfred flowers. It would be good to relieve Mildred's burden, or make Alfred happy, so we have reasons to do so. Intuitively, it is facts about *these individuals* which explain why we have these reasons.

These examples suggest that in many cases reasons covary with facts about neutral value, and this in turn suggests a strong conjecture:

SUFFICIENT EXISTENCE OF REASONS: If some fact constitutes or partly explains the fact that doing something would realise some state of affairs that is valuable, that fact is a reason for you to do that thing.³⁶

The fact that helping Mildred with her bag will relieve her burden is a reason to help her, since her burden's being relieved is a neutrally valuable state of affairs. The fact that giving Alfred flowers will please him is a reason to give him flowers, because his being pleased is valuable.

We're talking here about *neutral value*. This is a technical name for a familiar idea, namely that value is a monadic property of objects or states of affairs, and which admits of degrees. States of affairs consisting in things being good for people may well be neutrally valuable, in this sense.³⁷ Hereafter I'll refer to neutral value simply as 'value.'

This modest sufficiency claim – that facts about value are sufficient for there to be reasons – is a compelling motivation for value-based theories. Once we have decided that some state of affairs is valuable, who could deny that the fact that doing something would realise that state of affairs constitutes *a* reason to do it?

³⁶ I prefer the terminology of 'value' and 'disvalue' to that of 'good' and 'bad' since the former is broader (e.g. extends to aesthetic and prudential value; to some ears 'good' and 'bad' have a moralistic connotation I wish to avoid). However, as above, I'll occasionally use 'good' and 'bad' to avoid overly cumbersome prose; still, I'll be using them to pick out the property of *being valuable*.

³⁷ An alternative strategy abandons the idea of neutral value for a notion of relative value (cf. Michael Smith 2003 and Douglas Portmore 2001, 2003). For a state of affairs to have relative value is for that state of affairs to be valuable *relative to* someone or other. On this view, states of affairs are never valuable in and of themselves, or in abstraction from some or other point of view. I'll set this alternative aside. For worries about this approach, see Mark Schroeder 2007.

This is of course consistent with allowing that there might be overwhelming reasons not to do it.

The necessity claim – that a certain relation to facts about value is necessary for some fact to be a reason – would be more contentious. T.M. Scanlon, for instance, allows that there might be reasons that are not correlated with values, for instance reasons arising from promises and other commitments.³⁸ Strictly speaking we don't need to suppose that the necessity thesis is true for our purposes here. However the principle I will defend will make this additional necessity claim more plausible.

We also think that the *strength* or *weight* of these reasons varies with the value of these states of affairs. If Mildred could probably manage on her own, but with a little discomfort, then you have a fairly lightweight reason to help, perhaps counterbalanced by the reason not to damage her pride. If she looks as though she will collapse under the weight of the suitcase, you have a much stronger reason to help. Likewise with Alfred's happiness, and the child's suffering.

We can express this idea in the following principle:

VALUE-WEIGHT CONNECTION: When some reason is explained by some fact about value, the greater this value, the greater the weight of the reason, other things equal.

³⁸ Cf. Roger Crisp 2005, p.83.

I have argued for stronger versions of these two principles elsewhere.³⁹ It is important to note that these weaker principles allow not only for non-value-based reasons, but also for buck-passing accounts of value, since they don't mention anything about explanatory direction. For the purposes of this paper I'm going to assume that these principles are true. We're interested in a problem that arises within this way of thinking about ethics.

The problem arises when we try to accommodate the normative significance of special relationships, personal projects, and ethical and other commitments. These are traditionally considered to be the stronghold of the deontologist and the virtue ethicist. I'll now present three cases that we'll be discussing throughout. These cases present a deep problem for consequentialism. I'll shortly argue that the principle I'm going to defend not only gets the right results in these cases, but it gets these results for the right reasons. It is a further advantage of my principle – perhaps even an advantage over deontological theories – that it explains these different kinds of cases systematically.

SPECIAL RELATIONSHIPS: Angela and Becca both love their mothers. Angela's mother is a beneficent soul, always doing her best to improve the lives of those

³⁹ In chapter one. These principles are all formulated as necessary and sufficient conditions, rather than as explanatory theses or analyses. It strikes me as more plausible that the facts about value and realisation relations *explain* facts about reasons, rather than the facts about reasons explaining facts about value, as buck-passers such as T.M. Scanlon would have it. I will often talk as though this is the case, but these principles and the Harmony Principle I'll be defending in a moment are consistent with buck-passing theories of value, with appropriate modifications.

around her. Becca's mother is rather self-centered, and never really thinks about or does much for others. Alas, the two mothers are stuck in a burning building. There is only one oxygen mask, so that only one daughter can save her mother. Becca is holding the mask (she came into its possession by chance). Becca can either save her mother or enable Angela to save Angela's slightly more beneficent mother.⁴⁰

ETHICAL IDEALS: An evil terrorist insists that unless Bertrand kills an innocent person, two other people just like Bertrand will themselves be forced to kill innocent people. (We can assume that the person Bertrand would kill would be one of these two. If the others do not do this something terrible would happen, like the detonation of a doomsday device.) Bertrand can either kill an innocent or ensure that two others will kill innocents.⁴¹

VOLUNTARY COMMITMENTS: Peter, Matthew, and Mark all promised different colleagues that they would buy them a ticket to the show. As it turns out, Peter knows that if he keeps his promise to get his colleague a ticket, the others will be unable to keep their promises. All three promises were made in similar circumstances; all three gentlemen left it just a little bit too late to keep their promises, so the fact that there are no tickets left would not be an excuse. Peter can either keep his promise or enable the two others to keep their promises.

⁴⁰ This is a modified version of the famous case in William Godwin 1793.

⁴¹ This is a familiar kind of case; see e.g. Judith Jarvis Thomson 1997, p.274. It is slightly different from Bernard Williams's 1973 case.

These cases have the following features by stipulation. An agent has just two alternatives actions available to her: A and B. A will realise more value than B. There is no kind of value realised by B that is not realised by A, nor any kind of value realised by B in a larger amount than the same kind of value realised by A. The agent is in some intuitive sense *partial* to the slightly less valuable action B, and not to action A.

Our question is whether our protagonists have weightier reasons to do the thing they are partial towards or to do the thing that would bring about more value overall.

And here is our datum. *It is extremely plausible* that in each case the protagonist has more reason to do the thing he or she is partial towards. Becca has more reason to save her mother than to allow Angela to save Angela's mother. Bertrand has more reason not to kill an innocent than to prevent two others from killing innocents. Peter has more reason to keep his promise than to enable Matthew and Mark to keep their promises.

So what is the puzzle? The puzzle is that our two theses concerning the existence and weight of reasons pull us towards a familiar stronger thesis about the weights of reasons, which these cases suggest is false. The thesis is this:

REASONS CONSEQUENTIALISM: the weight of your reason to do something is directly proportional to the amount of value that explains the reason in question.

This is a reasons analogue of act consequentialism. This thesis seems to be a natural development of the idea that facts about reasons are to be explained in terms of facts about value. It is a powerful extension of the Value-Weight Connection: it provides necessary and sufficient conditions for some reason's having some particular weight. But Reasons Consequentialism gets the wrong result in these three cases, for the cases are designed in such a way that more value will be realised by the alternative to the partial action. Since we want to get the right result – indeed more, we want to *vindicate* the intuitions expressed in our three cases – we must reject Reasons Consequentialism and look for a better alternative, which is consistent with our two principles and yet vindicates this datum about our three cases.

In the following section I will introduce and motivate my own principle. In the final section I will compare this principle favourably with reasons consequentialism, indirect consequentialism, and global consequentialism. I conclude with a few further advantages of this new way of thinking about things.

SECTION TWO: THE POSITIVE PROPOSAL

I offer the:

HARMONY PRINCIPLE: The weight of a reason for you to do something is an increasing function of the amount of value that explains the reason

in question, and the value of your having a pro-attitude towards this value.

It will become clear as we proceed why I think the principle deserves this name. This principle maintains that the weights of reasons respect a certain harmony between the value of outcomes and the value of attitudes to those outcomes – and indirectly with the value of a certain kind of life lived. It also purports to accommodate attractive features of consequentialism, deontology, and virtue ethics. I'll argue that this principle is attractive and that it can solve our puzzle about partiality.

According to the Harmony Principle, the weight of a reason is affected by two different factors: the value that would be realised by doing the thing, and the value of your having a pro-attitude towards it.⁴² This principle entails two weaker principles. The first maintains that, other things equal, the greater the amount of value that explains some reason, the greater the weight of that reason. This is just the Value-Weight Connection; as we have seen it is very plausible.

⁴² There may be other factors as well. It is plausible that weight varies with amount of value as well as with the likelihood that the action in question will realise that outcome. Suppose you have a one in ten objective chance of making Alfred 10 units worth of happy, and a five in ten objective chance of making Mildred 10 units worth of happy, and you have to choose. They are both strangers to you, and we assume these ten units would have the same beneficial impact on their lives. Were the chances the same, perhaps you ought to toss a coin. But with these odds, clearly you have more reason to try to help Mildred rather than Alfred. It is similarly plausible that we understand temporal discounting in terms of a value-based weight-determining function with an additional factor. This suggests that an approach similar to the one to be defended here might help a value-based theorist of normativity also to accommodate *temporal* partiality. Ralph Wedgwood adds a factor for *agential involvement*, see his 2007, p.15.

The second principle maintains that the greater the value of having a pro-attitude towards the value that explains the reason in question, the greater the weight of that reason. We'll do some work to make this more precise, but it should be initially quite attractive. It is plausible, in each of our three cases, that it would be better for the agents to want to perform the partial action more than the alternative. It is plausibly better for Becca to want to save her mother, for Peter to want to keep his promise, for Bertrand to want not to kill more than he wants to save two lives. These are fairly deep and important facts about how it is good for these individuals to be oriented.

There are three claims here for which I will now go on to provide some support. Firstly that the relevant attitudes are valuable. Secondly that the value of these attitudes is related to certain facts about how it is good for these individuals to live. And thirdly that the values of these attitude covary in the appropriate ways with the weights of the reasons in question. I'll motivate these in turn.

THE RELEVANT ATTITUDES ARE VALUABLE

Start with Becca, who had to choose between saving her mother and enabling Angela to save Angela's more beneficent mother. It is good that Becca wants her own mother to be saved more strongly she wants some stranger's more beneficent mother to be saved. Similarly, it is good that she is more averse to her own mother dying in the fire than to someone else's mother dying. It is

certainly good that she cares about both. But it would be almost monstrous if she were indifferent to whether her mother or a stranger perished in the fire.

Obviously some relationships are better or worse than others. Perhaps there is no value in having a special attachment to a loveless and abusive partner. Importantly the Harmony Principle maintains that it is the *value* of the attitudes that plays the normative role, not the attitudes themselves. Those who desperately want to financially or emotionally support their abusive partners may not have weighty reasons to do so. By contrast those who have no desire to extend a special concern to their loving friends or family may nevertheless have weighty reasons to do so.⁴³

Now let's think about ethical commitments. Remember Raskolnikov's torment in *Crime and Punishment*. Or Jesse Pinkman's in *Breaking Bad*. It's a terrible thing for a person of good conscience to kill someone, and your average citizen has a very strong desire not to do so, and strong negative emotions afterwards if they do. This is plausibly a very good way for people to be. Of course this might vary with a person's situation. On the assumption that it will sometimes be good to have soldiers who are trained to kill people, it is better that they do not suffer too much when they return from duty. But this is clearly not an ordinary case.

⁴³ What kind of attitudes would it be good for a developed-world hermit to have? Well, first regarding actions: we know that she should pay her taxes, contribute a decent proportion of her income to charity, treat passing strangers kindly – and it would be good for her to want to do all these things for the right reasons. But it is not obviously good for her to be partial in her attitudes towards anyone, neither would it obviously be good for her to want to engage in social activities, projects, etc (little league baseball on Saturday mornings). Though still she might be partial to, say, detective novels, or Wagner.

The contrasting case is expressed in the idea that ‘virtue is its own reward.’ It is good, at least, for an individual to be happy to do a good thing for someone, to be pleased afterwards – not self-satisfied, but positively oriented towards doing good things. Alas, virtue is not always in fact its own reward. Some people do not have the right orientation towards the performance of good acts. The sour puss may find no reward in beneficence. What *is* true is that, for most people, it is good for them to be such that virtue to be its own reward. Why is that? Well, by and large, it is good for people to have pro-attitudes towards states of affairs that are valuable, and moreover towards states of affairs consisting in states of affairs being valuable, i.e. towards value *as such*.⁴⁴

We can also consider voluntary commitments. It is good if individuals, in general, are disposed to keep their word, to do what they say, to speak sincerely, and so forth. In addition to the value of these dispositions, plausibly it is good to have other associated pro-attitudes: to feel some grief about having to lie to a friend in order to ensure a much greater benefit for her, to take pride in the fact that ‘your word is your bond.’ So again, plausibly, it is good if Peter is more strongly disposed towards keeping his own promise, than towards enabling others to keep theirs.

⁴⁴ Claims of this sort have received a hearty philosophical defence in Thomas Hurka’s *Virtue, Vice, & Value*. He defends a principle similar to the following (2003, p.13): if S is some finally valuable state of affairs, then having a pro-attitude towards S is also finally valuable. This principle is very attractive. However it says nothing about the *amounts* of value or about *strengths* of pro-attitude. I think it is important to see that these can vary systematically. The cases I have in mind more closely accord with the following principle: if S and S* are two valuable states of affairs distinguished only by the fact that agent A is *partial* towards S and not towards S*, then it is better for A to have stronger pro-attitudes towards S than S*.

These are substantive axiological claims. I submit that these claims are plausible, and I'll proceed to argue that we can appeal to the plausibility of these claims in the explanation of our responses to our three cases. We are, in effect, passing the buck in our account of partially weighted reasons to these claims about the value of partial attitudes.

WHY IS IT GOOD TO HAVE PARTIAL ATTITUDES?

There is a venerable consequentialist tradition according to which the value of attitudes is just their instrumental value. Jeremy Bentham, for instance, holds this view.⁴⁵ But this is not the kind of value that we are interested in. The instrumental value of an attitude is neither sufficient nor necessary to play the role in question in the weight-determining function. Against sufficiency, consider a variation on a 'wrong kind of reasons' case, i.e. one in which some great incentive or deterrent for an attitude is in place. Suppose the gunman will shoot you unless you want to bring about some terrible outcome.⁴⁶ Intuitively there would be value in your having this attitude, but this value would not affect the weights of any reasons for you to bring about the terrible outcome. Instrumental value does not seem necessary either. To paraphrase Jonathan Way, it is plausibly good 'to want people you have never met to be happy, even

⁴⁵ Jeremy Bentham 1789, p.102.

⁴⁶ Compare Pettit & Smith, 2000, p.132.

if your wanting this has no effect on whether they are happy, or any other good effects.’⁴⁷

One natural alternative is to restrict our attention to attitudes that are finally valuable. The final value of a state of affairs is the value it has for its own sake, or its non-instrumental value.⁴⁸ It is not possible to concoct standard ‘wrong kinds of reasons’ cases in which the attitude has *final* value. Moreover it is very plausible that in many cases the attitudes in question will be finally valuable. This seems true of at least two of the three cases under consideration: SPECIAL RELATIONSHIPS (and many personal projects) and ETHICAL IDEALS.

I’m not so sure about VOLUNTARY COMMITMENTS. Plausibly it is good to be firmly disposed towards keeping promises - but not because that disposition is finally valuable. Plausibly this disposition is valuable because it is in a certain way instrumental to respecting others, being concerned not to harm others, being trustworthy, and so on. Plausibly we can say the same of a disposition towards punctuality or truthfulness. In fact, many ‘executive virtues’ will also be like this, for instance the virtue of decisiveness, or perhaps attentiveness. We need some way to distinguish those attitudes which, though not finally valuable, nevertheless intuitively play the relevant role in the weight-

⁴⁷ 2013, p.20.

⁴⁸ This is often called *intrinsic* value. This label is unfortunate since it forecloses the theoretical possibility that there will be cases of final extrinsic value (e.g. rare stamps and wedding rings). I take no stand on whether there are such cases. Notice that the attitude itself is a complex state of affairs, consisting of a subject, a relation, and a state of affairs, e.g. <Jill’s being pleased that Alfred is happy>. Importantly we are talking about the value of the attitude itself, rather than just the value of the object of the attitude.

determining function, from those attitudes which are instrumentally valuable and which clearly do not.

At this point we can most profitably appeal to attitudes the value of which is partly constitutive of living well. This axiological notion enables us most effectively to distinguish those valuable attitudes which do from those that do not intuitively play the role in the weight-determining function. Partial attitudes are good because having them is partly constitutive of living well, or flourishing. Another mark of this is that most people would be in an important sense *alienated* if they did not have a decent range of these kinds of partial attitudes towards projects, friends, ideals, etc.

Why is it that partial attitudes partly constitute living well? Here I find it helpful to recall a remark from Wittgenstein's *Philosophical Investigations*:

“We have got onto slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk. We want to walk: so we need friction. Back to the rough ground!”⁴⁹

If we were all perfect agents, able to interact with any given person equally well, able to engage in any given project equally well, able to perform any act in any situation without compunction or delay – and if, moreover, we had the resources to engage in any number of relationships, projects, and careful ethical

⁴⁹ Ludwig Wittgenstein, 1953, Pt 1, §107.

discriminations – if we were all like that, *possibly* it would be good if we had impartial attitudes. But clearly we are not such people. We find ourselves with very different predilections, tastes, and talents, and different historical relationships with people: with our family, our nation, the person we sat next to in primary school. All of this means that we are bound to form more valuable relationships with certain people than with others, more valuable commitments to some projects than others, and so forth.

These relationships in turn are constituted by certain personally-directed emotional and conative attitudes, and dispositions to have such attitudes. This will often consist in feeling sad when bad things happen to those one loves, and pleased when good things happen. What is good is to be *directly responsive* to events in these ways. This is part of what it is to be emotionally committed to someone.⁵⁰ We think it is *good* to be committed in this way. Plausibly it is good to have some emotional sensitivity to all sentient creatures like this. But also plausibly, it is good to have a *stronger* degree of emotional sensitivity to one's family and friends.

This appeal to the more general axiological notion of the good life is strictly additional to the Harmony principle. Nonetheless it plays a number of useful

⁵⁰ There isn't a good word for this in English. In the polyamorous community, the word "compersion" is sometimes used for a state like this. Perhaps this is what Saint Paul had in mind when he asked the Ephesians (5.21) to "be subject to one another." Perhaps also what Octavio Paz had in mind when he said in *Piedra De Sol* that "to love is to undress our names." Never mind that he says a little earlier that to love is to battle. Here are the surrounding lines:

el mundo cambia / si dos se miran y se reconocen, / amar es desnudarse de los nombres: / "dejame ser su puta," son palabras / de Eloisa, mas él cedió a las leyes, / la tomó por esposa y como premio / lo castraron después

roles in this account. It distinguishes those valuable attitudes which do from those which do not play the normative role in question. It provides additional and more general axiological support for our more particular axiological claims. Finally it provides an additional systematic basis for the appeal to valuable attitudes in the weight-determining function. It is rather plausible that facts about the good life play a fundamental role in the explanation of how weighty our reasons are, and, eventually, what we ought to do.

THE WEIGHTS OF REASONS COVARY WITH THE VALUES OF THE RELEVANT ATTITUDES

We shall now motivate the claim that the weights of reasons covary with the value of these attitudes in the manner specified. The way to argue for this is to compare cases in which the value of the outcome is fixed, and we vary the value of the agent having a pro-attitude towards that outcome to see whether our intuitions about weights track these changes.

Think again about Becca; imagine her relationship to her mother being increasingly less a part of what makes her life worth living. Perhaps her mother has been loveless, or absent, or abusive. Plausibly, still, Becca should save her before a stranger. But also, plausibly, the weight of her reason to do so diminishes along with the value of her relationship with her mother.

Take another case. Suppose that either some ecologically rich patch of rainforest has to be destroyed, or some chimpanzees have to lose their lives. Let the patch of rainforest, teeming with life, be marginally more valuable; plausibly Jane Goodall would still have more reason to fight to save the chimps.

Here's a case in which the value of having a strong emotional aversion to suffering affects the weights of reasons. If either you or a battle-hardened field nurse had to chop off someone's finger off to save their life, the nurse would have more reason to do it. Here's a case in which the value of someone's having a strong aversion to killing people affects the weights of reasons. If either you or James Bond has to kill an innocent person, to avert catastrophe, Bond would have more reason to do it.

What about this variation on this last case. Kevin has no aversion to killing people, though in fact he wouldn't except to avert catastrophe. Suppose he and Bond are in the aforementioned scenario, and let's imagine that the technology is such that either Kevin or Bond could do the deed with equal effectiveness. Finally stipulate that not only is Kevin not averse, but he wouldn't suffer any more than Bond if he had to kill the person in this case. Plausibly, it would be better if Kevin were averse to killing people, and plausibly it is not the case that it would be better if Bond were, but plausibly their reasons to kill the innocent person in this unfortunate case are of approximately equal weight. This would be a counterexample to the Harmony Principle. In response, I deny this claim about equally weighted reasons. The Harmony Principle appeals to facts about which attitudes are valuable, rather than which attitudes individuals actually

have. This has the great advantage of getting the right results in cases in which agents lack attitudes it would be better for them to have. Even if Becca did not feel compelled to save her mother, she would still have weighty reasons to do so. This result seems to me to hold in full generality. Suppose now either you or Kevin has to do the deed. You have additional reasons not to do it, since you will suffer greatly afterwards. Suppose that isn't the case; suppose you know you won't remember either way. Of course you'll still suffer as you decide to kill the person and actually do so. For these reasons you have more reason not to do it than Kevin. But the weightiness of one's reasons not to kill innocents has very little to do with the weight of your reason to avoid this psychological unpleasantness. It has much more to do with your reasons not to do something that it is good for you to be averse to doing. Notwithstanding reasons arising from your suffering, it seems plausible that you and Kevin have approximately equally weighted reasons to do and not to do the deed – and both of you greater reason not to do it than Bond.

Again, all this supports the idea that if you hold the value of the outcome fixed and vary the value of a pro-attitude towards the outcome, you will vary the weight of the agent's reason to bring about the outcome. It was already plausible that holding the value of the pro-attitude towards the outcome fixed and varying the value of the *outcome* also had the effect of varying the weight of the agent's reason to bring about the outcome. In both cases these are increasing functions – the greater the value, the weightier the reason. Hence we should think that something like the Harmony principle is true.

WHY THE HARMONY PRINCIPLE WORKS

So far I have attempted to motivate the claim that it is finally valuable for the protagonists in our three examples to be attitudinally partial towards the less neutrally valuable alternative. I argued that these axiological claims are individually plausible and admit of a more general explanation in terms of the life it is good to live, and that the absence of these attitudes leads to a certain kind of alienation. Let me first show how this enables us to get the intuitively right results about the weights of reasons, then I'll address why we should think the Harmony Principle is true.

We can insert some dummy numbers to model the cardinal values of the outcomes and attitudes in Peter's case, which will serve as a schema for the rest. You may recall that Peter was faced with the choice of keeping his promise, or breaking his promise and thereby enabling two others to keep theirs.

	Keeping Promise	Breaking Promise
Value of Outcome	5	10
Value of Attitude	9	3

We haven't said much about the nature of the function involved in the Harmony Principle, except to say that it is an increasing function of these two focal points. I think the Sufficient Existence of Reasons principle puts a further constraint on this function, but we needn't worry about those details for now. Let the function simply sum the quantities in its argument places. Then the weight of Peter's reason to keep his promise will be 14, and the weight of his reason to enable others to keep their promises will be 13. Hence, he has more reason to keep his promise. This is obviously very schematic. There are important and difficult questions about the epistemology and metaphysics of the function and issues about the determinacy and commensurability of amounts of values. I'm optimistic about the latter, and for now defer to Ruth Chang's excellent work on the topic.⁵¹ About the former it will be sufficient to point out that there is no reason to think that the Harmony Principle will encounter epistemological or metaphysical disadvantages that do not similarly afflict any of the familiar consequentialist principles.⁵²

Let me now say a word about how exactly this principle turns the trick of getting relative weights out of neutral values.

The plausibility of Reasons Consequentialism relies upon the fact that actions are in an important sense evaluatively *thin*. On a consequentialist way of thinking about the value of actions the fact that if I perform action A *you* will bring about some valuable state of affairs S, and if I perform B *I* will bring about

⁵¹ Cf. in particular Chang 1997.

⁵² For a little more on this topic, see footnote 24.

S – this distinction irrelevant to the action's normative status. The agent herself is not guaranteed to be part of any evaluated states of affairs, nor would her involvement have any special significance. We can think of actions as the merest intervention by whoever "happens to be near certain causal levers at a certain time" (as Bernard Williams puts it⁵³). This is a coherent and occasionally attractive way of thinking about the normative assessment of actions.

It is not possible to assess attitudes this way. As we have seen, attitudes always consist in a relation between an agent and a state of affairs. The agent herself *cannot but* stand in the first place of this relation. States of affairs consisting in some agent standing in an attitudinal relation to some other state of affairs are regular old states of affairs like any other, and admit of some neutral value or disvalue just like any other. However they are in one important and non-axiological sense agent-relative, simply because the agent cannot remove herself from the first relatum. It is metaphysically impossible for *me* to have *your* attitudes, or you mine. (There is a point of connection here with the fact that I cannot promise that you do something. At best I can promise that I will do my best to see to it that you do the thing.)

By appealing to the value of the agent herself having certain attitudes we inject an element of relativity into the states of affairs that are being evaluated. It is better for Becca to want her mother saved than for Becca to want to some other mother to be saved. Importantly what is being assessed is a pair of *Becca's* attitudes, not Becca's attitude to her mother and someone else's attitude to their

⁵³ In his 'Persons, Character, and Morality,' 1981, p.4.

own mother. In determining the weight of the reasons for *Becca* to do this or that, we partly appeal to facts about what it would be better for *Becca* to want or to care about. This, if you like, is how we turn the trick of getting relative weights from neutral values.

SECTION THREE: THE THEORETICAL ALTERNATIVES

REASONS CONSEQUENTIALISM

J.J.C. Smart maintains that “the [consequentialist] principle expresse[s] the attitude of generalised benevolence.”⁵⁴ But it is one thing to say one should love all creatures, and quite another to say that one should love them all *equally*. The latter is quite clearly a mistake. Impartiality in one’s actions is appropriate in an umpire, a teacher, and a judge. It is inappropriate in a parent, a lover, a friend, or, in general, a human being.

The Harmony principle has a few advantages over act consequentialisms of one sort or another. Consider first *consequentialising* consequentialism.⁵⁵ This is the strategy of adding to one’s axiology some (dis)value corresponding to any act-type which the deontologist thinks we have a duty (not) to perform: for instance, a final value in promise-keeping, a final disvalue in killing innocents, a final value in saving one’s mother. This strategy will get the right results for

⁵⁴ J.J.C. Smart, 1973, p.67.

⁵⁵ Cf. James Dreier 1996.

the Reasons Consequentialist in some cases, though to many it doesn't seem as though it gets these results for the right reason. But in any case this strategy does not get the right result in cases like our three, since they are designed to be such that *whatever* value realised by partial action, more would be realised by the alternative. The Harmony Principle gets the right result in these cases, and also has more to say about *why* it does so.

There is also an advantage over neutralising consequentialisms, i.e. those various attempts by neutral value consequentialists, such as Frank Jackson 1991, to explain partiality by appealing to the fact that one is epistemically or practically better placed to help one's friends and family, or that doing so will be better overall in the long run. This is intuitively the wrong explanation for why, for instance, one has more reason to save one's own mother from a burning building. Furthermore such accounts cannot explain the fact that one *would* have more reason to help one's friends and family, even if these more global consequences failed to obtain, i.e. even if doing so weren't in fact eventually productive of the best overall state of affairs.⁵⁶

GLOBAL CONSEQUENTIALISM

Consider the following view:

⁵⁶ A related advantage here is that this approach will avoid the worry for act consequentialists that one will have to be always looking over one's shoulder for greater value-creating opportunities. "A great concern to squeeze out the last drop of utility is likely to be a great impediment to the enjoyment of life." Adams, 1976, p.471.

GLOBAL CONSEQUENTIALISM: For any focal point (e.g. acts, rules, motives, etc.) F , and any set $x_1...x_n$ of available ways that focal point might be, one has most reason to have that x_i that would have the best consequences.

Crucially, each focal point is assessed separately, and hence the normative statuses of each focal point are completely independent. Hence according to Global Consequentialism you ought to have whichever attitude set would realise the best consequences; you ought to follow whichever decision procedure the following of which would realise the best consequences, perform that action the performance of which would realise the best consequences, and so on. This view is consistent, and it has some very high profile adherents.⁵⁷

There are two important problems with Global Consequentialism. Firstly it entails Act Consequentialism and hence, like Act Consequentialism, gets the wrong results in our three problematic cases.

Secondly it leads to odd cases of conflict between these different factors. Global Consequentialism tells us to perform the best action, have optimific attitudes, and to use optimific decision procedures. But then there will be situations in which Peter ought to be strongly motivated to keep his promise, ought to think

⁵⁷ See Peter Railton 1984; Fred Feldman 1986, 1993; Shelley Kagan 2000; Philip Pettit & Michael Smith 2000; Toby Ord (manuscript). See also Derek Parfit 1984, p. 24-28.

that promises ought never to be broken, but still he ought to break his promise. Thus Global Consequentialism leads to cases of 'normative conflict.'⁵⁸

By contrast, the Harmony Principle avoids unnecessary conflict. It is designed to get the right answers about what one ought to do in any situation, *given* all the facts about one's relationships and so on. The Harmony Principle provides us with a master principle that plays the role of balancing different higher-level normatively relevant factors.⁵⁹

An ethical theory should *take into account* the importance of honouring one's friendships, and the value of being directly responsive to one's friends' needs. The Harmony Principle gets the result that oftentimes one is doing exactly what one should be doing when one is helping one's friend rather than performing some other more optimific action. Moreover – with J.J.C. Smart now, and against Bernard Williams – in one's quieter moments one can conscientiously endorse one's dispositions to act exactly this way.

⁵⁸ Cf. Gerald Lang 2004; Jennie Louise 2006.

⁵⁹ We can compare this with Philip Pettit's 'red flag' approach (1991, 1997). In Pettit's view, one acts in accordance with one's motives (not: the motives it would be good if one had) until one notices a 'red flag.' His example is that one should want to help friends in distress until they ask you to move a body. But the problem with this approach is that, strictly speaking, there will be lots of cases in which one is not doing what one ought to do.

The 'red flag' approach is concerned with reasoning. Roughly speaking it advocates using a reasoning principle that gets results a bit like the Harmony Principle. But notice that its prescriptions for action conflict systematically with those of global consequentialism. For in many cases it will maintain that one ought to act in line with one's partial motivations even if this is not value-maximising, so long as it doesn't have a *very* detrimental effect on one's value production. On the other hand the red flag principle's prescriptions (so to speak) may well correlate roughly with those supported by the Harmony Principle. This suggests two further advantages of the Harmony Principle: on the one hand, it closes the gap between what one ought to do and how one ought to think about one ought to do (a gap which global consequentialism is forced to encourage) and on the other hand it provides a systematic account of whatever is attractive about the red flag approach to thinking about moral issues.

INDIRECT CONSEQUENTIALISM

An indirect consequentialism has the following form:

INDIRECT CONSEQUENTIALISM (SCHEMA): For two focal points F and G, and some set $x_1...x_n$ of available ways that F might be, and set $y_1...y_n$ that G might be, and some relation R relating ways that F might be with ways that G might be, one ought to have that y_i that stands in R to whichever of $x_1...x_n$ is consequentially best.

We are interested in a specific version of indirect consequentialism, according to which the normative status of acts is a function of the evaluative status of attitudes. Specifically, let *indirect consequentialism* be the principle that that one ought to perform that action, from among those available, which someone would perform if they were acting on that set of motives, of those available, the having of which would be consequentially best. This principle is attractive for several reasons. The first is its very indirectness. Lest it be objected that the Harmony Principle is too indirect or too complex, we can point out that this has not been considered a significant disadvantage of Indirect Consequentialisms.⁶⁰

⁶⁰ The Harmony Principle is a little more complex than Indirect Consequentialism, since it maintains that the normative status of one focal point obtains in virtue of the evaluations of two different focal points, but this doesn't seem very worrisome – especially not if, as I argue in the main text, it is plausible that the normative status of the one focal point does indeed covary with the evaluations of both. One might worry that in addition to mere ideological complexity there is also further ontological commitment. For according to the Harmony Principle the weight of reasons covaries with facts about the evaluations of these two focal points *together with the function itself*. Different theorists may advocate different functions, and these will have direct and dramatic substantive implications. More impartially inclined theorists will give

Secondly, Indirect Consequentialism will in many plausible situations get the result that one has most reason to perform some action that would realise less neutral value overall than some alternative, and hence it will be, in one important sense, non-consequentialist.⁶¹ The point is just to remind ourselves that we are familiar with value-based theories that are non-consequentialist in this sense. Thirdly, indirect consequentialism will more nearly get the right result in cases like Bond and the nurse.

However Indirect Consequentialism also gets some cases systematically wrong. To see the problem with this principle, imagine a case in which you have to do something really awful to save many lives. Perhaps you have to torture or kill someone, or bring shame upon your family. If the alternative is bad enough, it will plausibly be true that you ought to do the awful thing. But it is not at all plausible that it would be good that you *want* to do the thing, or feel *pleased* or *excited* about the fact that you will have to do this thing. Sometimes you have to be cruel to be kind – but it would be really bad if you *wanted* to be cruel. Perhaps sometime we will have to abandon loved ones in order to dedicate ourselves single-mindedly to an important political task. In such a case one should do the right thing, but plausibly it would be really bad to *want* to do so.

partial attitudes little importance; more partially inclined theorists will give partial attitudes great importance. I take this to be a theoretical strength of the approach. These alternatives should not be ruled out *analytically*, but rather substantively. What about the worry that Harmony Principle does not provide a strictly *value-based* account of weight, since it appeals in addition to this function? This may just be a terminological worry. But one might also argue that the substantive issue we are discussing may well be an *axiological* one, on the grounds that some deeper value such as justice or fairness explains the relative importance of outcomes and attitudes in the weight-determining function.

⁶¹ Krister Bykvist 2002, Michael Smith 2003, and Ralph Wedgwood 2009 all take this to be a sufficient condition for an account to be non-consequentialist. According to this taxonomy, the Harmony Principle is a non-consequentialist principle.

The other direction of the indirect consequentialist's biconditional also fails. Plausibly we have little or no reason to do something prompted by valuable motives or attitudes in a case in which absolutely no good will come of doing so. If Peter's loyalty in some other case would actually hurt his mother, plausibly in that case Peter would have no reason to do the loyal thing.

It is tempting to object to these counterexamples with an analogue of the objection that rule consequentialism collapses into act consequentialism: namely by arguing that it is best to have those attitudes the having of which will tend to make one perform acts with the best consequences. But this objection cannot be sustained. The evaluation of attitudes is a completely separate matter from the evaluation of the actions which these attitudes would incline one towards. Robert Adams discussed a case in which Jack is strongly attracted towards spending more time looking around at Chartres Cathedral than he ought to according to act consequentialism. He says that in order for the present objection to be sustained,

“...one would have to show that the benefits that justify Jack's motivation by motive-utilitarian standards also justify his spending time on the choir screen by act-utilitarian standards. But they do not. For they are not consequences of his spending time there, but independent consequences of something that caused, or manifested itself in, his spending time there. It is not that deciding to devote only a cursory inspection to the choir screen would have put him in the wrong frame of mind for enjoying the visit. It is rather that, being in the right

frame of mind for enjoying the visit, he could not bring himself to leave the choir screen as quickly as would have maximized utility.”⁶²

Robert Merrihew Adams is the quiet hero of this paper. His ‘Motive Utilitarianism’ in 1976 anticipated these objections to reasons consequentialism and global consequentialism. Adams’s paper primarily defends a direct consequentialism, which grounds the normative status of motives in an evaluative status of motives. This is slightly to the side of our main interest here.⁶³ But in one paragraph he sketches an account of the normative status of actions based on the evaluative status of motives. What is interesting about his account is that it deviates somewhat from indirect consequentialism as we have just stated it. Adams says “...it is plausible...to say that Jack is not acting wrongly in acting on the motivation that he has rightly cultivated in himself. But I think that is because it is plausible to depart from act utilitarianism at least so far as to allow the rightness or wrongness of Jack's action in this case to depend *partly* on the goodness or badness of his motive, and *not solely* on the utility of the act” (*op. cit.* p.474; italics added). Here we have the familiar response to act and global consequentialism. But notice the qualifiers: the normative status of the act depends only partly on the evaluation of the motive, and also partly on the evaluation of the consequences of the act.⁶⁴ If Adams could be convinced to restrict the normatively relevant attitudes to those which

⁶² Adams, 1976, p.472

⁶³ Some version of the Harmony Principle is applicable to reasons for attitudes, but a full discussion of this would take us beyond the scope of the current paper.

⁶⁴ Similarly, he leaves unanswered “the question whether a conscience of the most useful kind would be offended by some acts that maximize utility – particularly by some utility-maximizing violations of such rules as those against stealing and lying” (*op.cit.*).

are related in the right way to living well, perhaps he would be happy to accept some version of the Harmony Principle.

CONCLUSION

Let me conclude by pointing out a few further advantages of this principle.

First we have a few advantages over act consequentialisms of one sort or another. As we have seen, these views cannot get the right results in cases like our three. These consequentialist approaches also give the wrong explanation for *why*, for instance, one has more reason to save one's own mother from a burning building. Consequently such accounts cannot explain the fact that one *would* have more reason to help one's friends and family, even if these more global consequences failed to obtain, i.e. even if doing so weren't in fact eventually productive of the best overall state of affairs.

Secondly this helps us with Bernard Williams's famous 'one thought too many' charge against consequentialism. He objected that one ought to visit a friend in hospital *for her sake*, and not because doing so happens to maximise value. Of course, I agree. The Harmony Principle avoids this latter result for three main reasons. First, it is a principle concerned with reasons rather than oughts, so it is already in the business of giving particular states of affairs normative significance. Secondly, the Harmony Principle explains why one may well have most reason to see one's sick friend even if doing so would not be productive of

the most value overall. Thirdly, this principle thoroughly endorses the attitudinal orientation which compels one to go to see one's friend. Williams is right about that much. However I think he goes too far. A further advantage of the Harmony Principle is that it actually fully *justifies* one in going to see one's sick friend. It fully and systematically explains why this is the thing one has most reason to do. Therefore this principle also fits nicely with Aristotle's idea of a transition from natural virtue to full virtue, at which latter point one has reflectively *endorsed* one's partial attitudes.

Thirdly this view will help us to avoid Henry Sidgwick's famous 'dualism of practical reason.' He claimed, roughly, that there were two kinds of normative considerations – egoistic and moral. Moral considerations are universal and impartial. Egoistic considerations are extremely partial. Sidgwick worried that these kinds of normative considerations were incompatible – that they could not be weighed against each other. We shouldn't accept this. Neither egoistic nor moral reasons are impartial. They are explained in structurally similar ways, and they are commensurable, except perhaps at the margins – but that's everyone's problem. We should reject any deep metaphysical or theoretical distinction between egoistic reasons, partial reasons, or moral reasons.

The Harmony Principle purports to preserve the main motivation for consequentialisms, namely the Sufficient Existence of Reasons principle and the Value-Weight Connection. It systematically accommodates the main motivations for deontological views, arising from the normative significance of special relationships, personal projects, and ethical and voluntary

commitments. Finally, it also retains central advantages of virtue ethical views, since it gives special normative significance to valuable attitudes and their connection with the good life.

CHAPTER THREE: OUGHT, MOST REASON, AND SUPEREROGATION

ABSTRACT

An increasing number of contemporary ethical theorists are attracted to a reasons-first account of normativity, and hence to a reasons-based analysis of the all-thing-considered ought. The most natural such analysis maintains that one ought to do whatever there is most reason to do. However there is an important class of putative counterexamples to this analysis, which can all loosely be described as supererogatory actions. In such cases it appears that there are weighty reasons to do something, and in some cases these reasons seem weightier than the reasons favouring any alternatives, but it is not the case that one ought to do that thing. I argue that none of these cases constitutes a counterexample to the analysis of ought in terms of most reason.

There are two classes of cases to consider. One class involves alternatives with some positive deontic status: one is required, obligated, duty-bound to do the non-supererogatory thing. I argue that in such cases we need to weaken the principles relating these deontic notions to the all things considered ought. Sometimes you ought to do the thing that is required, and sometimes not.

Another class of cases involves alternatives which are all putatively permissible. In some of these cases the alternatives will be equally well supported by reasons; such cases pose no obstacle to our analysis. I argue that in others, we really ought to do whatever is favoured by the balance of reasons. I present two considerations in support of this claim: a no-relevant-difference argument, and some remarks about the relation between ought and the appropriateness of criticism.

The most interesting cases of this latter sort involve ‘imperfect duties.’ It is characteristic of imperfect duties that they do not require that one do anything *at a time*. Rather they require that one do *enough* of a certain kind of action over a certain period of time. I appeal to a notion from contemporary axiology, that of *contributory value*, arguing that we can use this notion to explain the reasons and their weights in such a way as to get the right results in these cases, without troubling the analysis of ought.

OUGHT, MOST REASON, AND SUPEREROGATION

An increasing number of theorists are attracted to a reasons-first account of normativity.⁶⁵ The idea is that the relation of *being a reason for* is the most fundamental normative notion, in terms of which – together with any relevant

⁶⁵ Cf. T.M. Scanlon 1996, Derek Parfit 2011, Joseph Raz 1999, Mark Schroeder 2007, John Skorupski 2010, etc.

non-normative notions – all other normative notions are to be explained. One important part of this project is giving an account of facts about the ‘all things considered ought,’ i.e. of what one ought to do in some situation given all the normatively relevant facts. This is the ought which is the end-point of first-personal deliberation, when one is considering all the facts and trying to work out what to do. Similarly, deliberation with friends about some difficult issue presupposes this ought. So does discussion with friends about what some third-party ought to do, given all the facts.⁶⁶ Importantly, all sorts of different normatively relevant considerations are relevant to this ought – prudential considerations, moral considerations, cultural considerations, aesthetic considerations, and so on. Some of these considerations favour or disfavour the alternatives on offer. These are reasons. Reasons are relations between a fact, an agent, and some action or attitude, such that the consideration counts in favour of or against the agent doing the thing in question. Reasons have weights: some count in favour more or less than others. These weights can somehow be aggregated to generate total nett weights favouring different alternative actions available in a situation. This is some function of the reasons in favour of that action and the reasons against. These total weights can be compared. Often some alternative will be supported by a greater nett weight of reasons than any others. This is what the agent has most reason to do in the situation.⁶⁷ Perhaps the most natural and attractive analysis of the all-things-considered ought, then, is the following:

⁶⁶ We’ll completely ignore the various debates about subjective versus objective oughts.

⁶⁷ This, at least, is my favoured conception of reasons, weights, and weighing. Various matters are controversial, for instance whether the fact that reasons have certain weights explains the fact that this reason is weightier than that, analogous to heights and *taller than*, or whether the

ANALYSIS OF OUGHT: For something to be what you ought to do all things considered is for that thing to be what you have most reason to do.

Some evidence for this analysis is provided by ordinary decision making, particularly in difficult cases. Suppose you have to decide whether to stay in your current job, or take a research post in some distant country, or leave academia to begin a career in politics. All sorts of considerations are relevant: the fact that the other job is better, in various respects; the fact that there are many more aesthetic and cultural attractions in the distant country; the fact that you would be living so far from your family; politically, you might have to weigh such considerations as the likelihood of your having some beneficial impact, or even getting elected in the first place; or the fact that your partner values her privacy. Faced with this predicament, you would run through some process plausibly describable as 'weighing' the considerations for and against each alternative. It is just as plausible that weighing is applicable in less difficult cases. Should you take the free tickets to the game, given that you have the afternoon free, that they'll otherwise go to waste, that you have been looking around frantically for tickets to this game for weeks? Of course you should. The reasons in favour easily outweigh the reasons against.

direction of explanation is the other way around. See Mark Schroeder 2007 for a defence of the later. Another important issue is whether reasons aggregate additively/quasi-additively, or in some more complicated way. See Selim Berker 2007 for the notion of quasi-additivity. I have argued for additivity elsewhere. This is currently a minority view. For arguments to the effect that the interactions between reasons are more complex, see Scanlon, 1998, p.51; Dancy, 2004a and 2000, p.132; Kagan, 1988. See also Raz 2012; Schroeder, 2006, p.3. Finally some argue that the very idea of aggregating reasons is mistaken. On this see the debate between Dancy 2004a, Berker, *op. cit.* and Lechler 2012.

Here is another reason to find this thesis attractive. Like any good analysis, it promises to minimize the number of unexplained notions in our ethical theory, by explaining one in terms of the other. Some theorists accept that ought and reasons are analytically related, but argue that reasons should be analysed in terms of oughts, rather than the other way around.⁶⁸ I find this alternative very unappealing. It is in the nature of an *all things considered* fact precisely to obtain in virtue of all the relevant facts in the situation. But the point for now is simply that ought-first theorists agree about the virtues of explaining one of these notions in terms of the other (their disagreement is about what explains what). The examples we shall be considering present *extensional* problems for these analyses, so my positive proposal will also be of assistance to oughts-first-ers.

A SCHEMA FOR PUTATIVE COUNTEREXAMPLES

I shall be defending this analysis of ought from an important class of putative counterexamples. The cases I'm interested in involve two *alternatives* (in the following sense: two available actions such that the agent can only either perform one or the other), with the following features:

1. X-ing would bring about more value than Y-ing, perhaps considerably more.
2. X-ing lacks some positive deontic property: for instance X-ing does not have the property of being required.

⁶⁸ Cf. Broome 2013, Kearns & Star 2009.

3. Neither X-ing nor Y-ing have any negative deontic properties. If you like: both are permissible.
4. Plausibly, it is not the case that one ought to X.

Here's an example of such a case, from Mark Timmons and Terry Horgan (2010):

BASEBALL: Olivia has just moved into a new neighbourhood. She learns that one of her neighbours, Mary, loved going to baseball games with her husband, who has recently passed away. Mary no longer goes to the games since she has no-one to go with. Olivia has no particular interest in baseball. However instead of spending the afternoon unpacking boxes she could take Mary to a game to keep her company.

We can stipulate that going to the baseball game would be better than not, since Mary has time to kill. This isn't yet problematic for our ANALYSIS OF OUGHT, since we haven't yet said anything about the relation between value and the weights of reasons. We need to add the following plausible principle:

VALUE-WEIGHT CONNECTION: When some reason is explained by some fact about value, the greater this value, the greater the weight of the reason, other things equal.

In fact all we need is that *whatever* it is that makes supererogatory claims *worth doing* in some respect, that feature provides reasons, in proportion to the degree

to which the thing is worth doing. Value is the most natural candidate for this role. So for instance, there is a less weighty reason for Mary to take Olivia to the game if she'll only be mildly amused than if it would be the most exciting event of her year. Or to take another example, the more suffering one can alleviate, or pleasure one can bring, by volunteering at a soup kitchen on Thanksgiving, the weightier a reason one has to do so. This thesis is rather weak, and is consistent with any number of other ways of explaining weight.⁶⁹ For the purposes of this discussion we are going to assume that this principle is true.

Now, so long as there is nothing else to be said for Mary emptying her boxes that might outweigh her reasons to take Olivia to the game, it looks as though Mary has more reason to take Olivia to the game than to unpack her boxes. It is still plausible that Mary isn't required to take Olivia to the baseball game, though it would clearly be permissible to do so. But also, plausibly, it is not the case that Mary *ought* to take Olivia to the game. Hence, we have a putative counterexample to ANALYSIS OF OUGHT, and schema for generating plenty more.

The class of cases which fits this schema is broader than the class of supererogatory actions. Some cases of 'enticing reasons' will fit this schema as well.⁷⁰ Still, I'll refer to these as 'supererogatory' cases for convenience.

⁶⁹ I defended a much stronger version of this thesis in chapter one, namely that we can analyse weight in terms of various facts about neutral values. This analysis entails the relevant biconditional thesis. We only need the sufficiency claim here, that facts about values are sufficient to explain weight, and not the more controversial necessity claim, that only facts about values explain weight. We leave open the possibility that, e.g., facts about promises unrelated to facts about values explain weight.

⁷⁰ Cf. Jonathan Dancy 2004b.

I shall argue that ANALYSIS OF OUGHT is not threatened by any instances of this schema. In some cases we ought to do the thing that is intuitively supererogatory. In some cases we ought to do something other than the supererogatory thing. In all cases the facts about the all things considered ought correlate with the facts about what there is most reason to do.⁷¹ To take the example of Mary and Olivia: once we make it clear that Mary really doesn't have anything better to do with her time, and that taking Olivia would make her extremely happy, it is rather plausible that Mary ought to take her to the game.

After a quick word on ideological matters, this paper proceeds in three further sections. In section two we'll consider cases in which the alternative to the supererogatory act is *required*. We'll consider cases in which these requirements are putatively decisive and cases in which they are merely pro tanto. In section three we'll consider cases in which both alternatives are permissible, and neither is required. Here we'll consider Singer-style cases, cases involving reparation, forgiveness, and favours, and imperfect duties. In the short section four we'll provide some further motivation for the ANALYSIS OF OUGHT by considering the relation between the all things considered ought and the appropriateness of criticism.

⁷¹ That is, in fact, all I'll be doing here, namely replying to some *extensional* putative counterexamples. Strictly speaking this paper is neutral about the question whether facts about the all things considered ought explain facts about most reason or the other way around.

SECTION ONE: MATTERS IDEOLOGICAL

J.O. Urmson's 1958 article 'Saints and Heroes' kick-started contemporary discussions about the supererogatory. Urmson was objecting to an influential tradition in ethical theory according to which the tripartite distinction between prohibitions, permissions, and requirements is sufficient to draw all the substantive distinctions we need in normative ethics. His goal was to add a fourth category to this tripartite distinction to account for actions which are not required or prohibited, but permissible, and which also have some further properties which make them 'saintly' or 'heroic.'

It is important to see that there is a way of doing ethics according to which these four notions – or really three, if we define prohibitions in terms of requirements not to do things – are sufficient for drawing all the distinctions between cases that we need. I think this approach is inferior to one which includes or consists in the ideology of reasons and the all things considered ought.⁷² But here I simply note that these traditionalists have no specific quarrel with our ANALYSIS OF OUGHT, for they say nothing that could conflict with it. Compare: no thesis about the relations between feng shui and one's qi affect theses about the relation between the superego and the id. We can feel free to ignore this traditional approach for now. Henceforth we will restrict our

⁷² The traditional approach simply cannot draw as many distinctions as we need. Its on-or-off threshold notions inaccurately represent the complex, graded structure of practical normativity.

attention to interlocutors who attempt to use the ideology of requirements and permissions alongside the ideology of reasons and the all things considered ought.

Matters are complicated by the fact that we have various normative domains – the moral, aesthetic, prudential, etc. – which interact in complex ways.⁷³ Then we have the more general domain of the practical, which includes all normative considerations relevant to what to do, what attitudes to have, and so on. It is so far an open question whether our various bits and pieces of ideology apply in these other domains. In this sense there might be moral requirements, prudential permissions, aesthetic reasons, and so forth. We can assume that some consideration in one of these domains is a reason only if it is a practical reason, that is, only if it is a consideration that favours someone actually doing something or being a certain way. However it is open question, which we shall be discussing shortly, what relation obtains between the fact that one all things considered *morally* ought to do something, and what one *practically* ought to do all-things-considered. Our ANALYSIS OF OUGHT is concerned specifically with practical reasons and the all things considered practical ought. Strictly speaking ‘practical reasons’ is pleonastic; we’ll henceforth refer to these just as ‘reasons.’ To save words, we’ll drop the words ‘all things considered’ from ‘all things considered ought’ and the word ‘practical’ from ‘practical ought.’

⁷³ We will not consider epistemic normativity here.

SECTION TWO: REQUIREMENTS

We shall start by considering cases in which the alternative to the supererogatory act is required. A *requirement* is a deontic notion in any domain that is local and decisive.⁷⁴ By *local*, I mean that the fact that some act is required is grounded by some relatively local facts about that act – for instance that it is a promise-keeping, or a violation of duty, or an instance of making someone happy. By *decisive*, I mean that the following principle holds: if some act is (the only act) required in some domain then one ought, all things considered, according to that domain, to perform that act.⁷⁵ Moral requirements entail moral oughts; practical requirements entail practical oughts, and so forth.

Such notions present the following problem for the ANALYSIS OF OUGHT. We have our two alternatives, X and Y. Y is supererogatory: permissible, excellent, and not required. Y's excellence together with VALUE-WEIGHT CONNECTION ensures that there are weighty reasons to Y. X is required. We're assuming that there is also some fact about what one ought to do all things considered. Since requirements are decisive, if one is required in some domain to X then one ought to X according to that domain. All of this is consistent with there being no interesting relations correlating requirements and reasons. Now consider two further conditions. The first condition is that the nett weight of reasons to

⁷⁴ Sometimes people use the word 'requirement' to refer to an all things considered normative notion, in particular when talking about 'moral requirements.' For an example see Karl Schafer, (ms). Unfortunately I'm going to have to be somewhat stipulative with my terms since there are no sufficiently precise conventions. I will stay as close to ordinary usage as I can.

⁷⁵ We can assume, what will be irrelevant to our purposes, that if one is required to perform alternative (hence incompatible) actions then one ought to perform some one of them.

do whatever is required is less than the nett weight of reasons to do the unrequired alternative. Secondly the requirement in question must be decisive in the *practical* domain (as, for instance, requirements of etiquette or fashion are not). If these conditions are met then it would be the case that one ought to do something that is not supported by the greatest nett weight of reasons, and we would have a counterexample to the ANALYSIS OF OUGHT.

We will focus on moral requirements, since *prima facie*, and as a matter of the history of the subject, it is most plausible that these are practically decisive. One kind of traditional deontologist will reject any necessary correlation between moral requirements and reasons. She will say that the moral requirement to keep one's promises, not to lie, to respect one's elders, etc. obtains irrespective of whether in each situation there are weighty reasons to do whatever is required. She will insist, for example, that there is a requirement to keep a promise to do something onerous and silly for someone who has forgotten and won't remember and who wouldn't care anyway. On such a view the first condition is clearly met: we have requirements without weighty reasons. But, in spite of her protestations, it is implausible that in such cases the second condition is met. In such a case it is implausible that one *ought* to keep this promise rather than doing something considerably more important.⁷⁶

⁷⁶ It doesn't matter much why this condition is not met, whether because these are moral requirements and moral requirements are not practically decisive, or because these are not really requirements, or because these are practical not moral requirements, and practical requirements are not practically decisive, or even because these are exceptions to a *ceteris paribus* principle that requirements are practically decisive.

There are other more challenging types of cases, in which requirements are correlated with reasons, which tend to be rather weighty. The deontologist will add that at least some requirements are practically decisive. Hence she will insist that even if some non-required alternative were supported by weightier reasons one ought to do whatever is required. Let's consider the case which most strongly supports this way of thinking. An unfortunate doctor's alternatives are whether or not to quietly kill some old man who happens to be lying on a bed in the hospital corridor, in order to use his organs to save five young people with tremendous prospects. Our interlocutor will perhaps concede that in some such case there is more reason to kill the old man than not. But she insists that there is a moral requirement not to, and hence that one ought not do so.⁷⁷

It is plausible that one is morally required not to kill the old man. It is less clear whether one practically ought not do it, all things considered. If the old man has no family or friends, and he is going to live the rest of his short life in pain and dereliction, and killing him is the only way to save our budding teachers and trombone players and government employees with strong socialist sympathies – and no-one will find out, and it won't happen again, and so forth, then maybe that is what the doctor ought to do. But notice what we are doing here. We are playing around with the various reasons in favour of and against our alternatives. Doubtless there are weighty reasons not to kill an innocent man. This is a substantive datum that any theory of reasons will need to accommodate. But in some cases these reasons will be outweighed. It is

⁷⁷ This is an ought not case rather than an ought case, but similar considerations apply.

possible that in such a case one ought to kill the old man and save the young five. This may be so *even if* one is morally required not to kill the old man, and, indeed, even if one *morally ought* not to kill the old man. This is just to deny that moral requirements and the moral ought are always practically decisive.

The problem for this approach arises from the fact that their specific combination of features renders them substantively inapplicable. Requirements are local. Whether something is required (or required not to be done) is fully grounded in local facts about the action in questions: it is a killing of a man, the keeping of a promise the preserving of basic liberties. Hence requirements are in a certain way *modally robust*. We can restrict our attention to practical requirements; similar remarks will apply to requirements in other domains. Since practical requirements are decisive it follows that tokens of these act-types are such that one ought to perform them *in any situation*. Hence one ought to perform any such act no matter how weak the reasons in its favour and strong the reasons against. But, alas, for just about any putatively required act-type a sufficiently imaginative ethicist will concoct a situation, probably a rather gruesome situation, in which you ought to do some alternative, or for *required not* cases, in which you ought to perform the action in question. Even our old man really ought to get killed off once he gets gnarly enough, and the budding trombonists sufficiently plentiful. This strongly suggests either that these acts are not required, or, more plausibly, that requirements are not

practically *decisive*, they are just really rather weighty. So these cases present no problem for our analysis.⁷⁸

This ends our discussion of local decisive notions. Similar points apply *mutatis mutandis* to local pro tanto notions, such as moral obligations and perfect duties. Since these are pro tanto notions they are already in the business of being weighed against other considerations. Sometimes obligations are outweighed by other obligations, as when one promises one friend to meet for coffee and another to attend her graduation and one cannot do both. Sometimes obligations are outweighed by other moral considerations, as when one fails to fulfil one's professional duties in order to save a drowning child. Sometimes obligations are outweighed by prudential considerations, as when one misses a meeting with a student in order to have an emergency heart bypass.⁷⁹

⁷⁸ What about a fall-back position, that moral requirements are decisive unless defeated by the fact the counterbalancing reasons are weightier by some threshold than the reasons favouring the required thing. Suppose we grant our interlocutor this notion of a threshold and the notion of defeat. There is some precedent for this position, for instance some philosophers think that one is no longer required to keep a promise if the consequences of doing so would be sufficiently awful. Then we pick some case in which the nett weight of reasons to do the required thing are less weighty than the reasons to do some non-required thing, but we stipulate that the difference between the nett weights of these reasons is less than the threshold amount, so that the requirement is not thereby defeated. In such a case, our interlocutor will insist, one ought to do the required thing, though there is less reason to do so. This is an important challenge. Unfortunately I do not think this worry can be fully discharged here. In order to assess the plausibility of this proposal, our interlocutor needs to bring us her theories of reasons and requirements, and to spell out some cases in some detail. I cannot think of a case in which it does not seem either that the weight of reasons supports the thing required, or that that one ought to do the other thing. Some additional reasons for scepticism about this fall-back proposal will emerge from our discussion of permissibility.

An alternative fall-back position is to provide a characterisation of requiring *reasons*, and to distinguish them from other, *justifying* reasons. See Joshua Gert 2007.

⁷⁹ Similar remarks will apply to prudential obligations. Compare two situations. In the first, one has to decide whether to incur some trivial inconvenience in order to save many others from intense suffering. In the second, one has to decide whether to make some move which will be extremely advantageous to one's career (perhaps we stipulate that one is not a jewellery thief or a hedge-fund manager), but which will inflict some trivial inconvenience on others. Plausibly in the first case one ought to save the others, and in the second case one ought to advantage one's career. It is easy to model such considerations in terms of differentially

Similar remarks will also apply to moral praiseworthiness and moral blameworthiness. Here I am assuming that moral praise is appropriate in cases in which one does something to benefit another at some considerable cost to oneself, or which requires considerable effort.⁸⁰ (We'll consider a more general kind of criticism later.) Sometimes it is not the case that one ought to do the thing that would be most morally praiseworthy. Perhaps your alternatives are to study for your exam or help some feral cat escape from a muddy bog. It would be most morally praiseworthy to save the cat, but if your exam is important enough (perhaps you are sitting the bar tomorrow), you ought to study.⁸¹ We don't even need to appeal to prudential cases. A teacher has more reason to spend her day off preparing classes than to volunteer at an unpleasant and understaffed soup kitchen. Still, it is plausible that working in the soup kitchen would be more morally praiseworthy. What about morally blaming someone for doing what they ought to have done? Perhaps Gaugin is an example of this – or if you prefer, Schmaugin: a character psychologically unable simultaneously to care for his family and make art, and so incredibly talented that he ought to abandon his family to paint. Plausibly he is still morally blameworthy for abandoning his family.

Let's end this section by looking more closely at the 'moral ought.' As an all things considered notion, the moral ought obtains in virtue of all the morally

weighted reasons. These reasons have different weights, and no such case presents an obvious counterexample to ANALYSIS OF OUGHT.

⁸⁰ *Mutatis mutandis* for blame. I call this 'moral praise and blame' to distinguish it from what you might call 'rational criticism,' which we'll discuss later.

⁸¹ Doug Portmore has some examples like this; cf. his 2008.

relevant facts, together with the fact that those are all the morally relevant facts. Now consider the thesis that you morally ought to do whatever you have most reason to do. This is counterexample-able in both directions. Sometimes I have most reason to brush my teeth or take a bath, but it is never (or rarely) the case that I morally ought to brush my teeth. Perhaps I morally ought to stay with my sick mother, but I have most reason to go to the office to close a deal. It is a tricky and contentious question which considerations are morally relevant. It seems plausible, at least, that some non-moral reasons will be morally relevant in this sense.⁸² It is an advantage of the view that I propose that this question doesn't even need to be addressed. So long as there are some normatively relevant considerations that are not morally relevant considerations, we can construct a situation in which weighty considerations of this sort are at stake and only lightweight considerations of the moral sort are at stake. In such a case, the morally relevant considerations will be outweighed, and so it will be false that you morally ought to do whatever you have most reason to do.

SECTION THREE: PERMISSIBILITY

Another key feature of Urmson's characterisation of the supererogatory is that both the supererogatory act and the alternative are *permissible*. The notion of permissibility presents problems for the ANALYSIS OF OUGHT as follows.

⁸² Neither is it straightforward to determine which facts are relevant to the moral ought. Suppose, what is itself not a simple matter, we can distinguish moral reasons from others, perhaps by notion of the "location" of the value that explains the reason, or by some fact about the duties, obligations, expectations, and so on. Still, as Joel Feinberg 1961 and Elizabeth Harman (ms) have argued, it doesn't seem as though only moral reasons are relevant to the moral ought. Prudential reasons are also relevant.

Suppose that X-ing and Y-ing are both permissible. Now suppose that there is more reason to X than Y. Now we assume two further plausible principles to get the intuitive problem going. Firstly we have a principle relating permissibility and the all-things-considered ought: if it is permissible to Y, then it is not the case that you ought not Y. Secondly we have a principle concerning ought and ought not: if you ought to X, and Y is an action incompatible with X-ing, then you ought not to Y.⁸³ This second principle is a little stronger, but we'll grant it to get the objection going. These principles together with our assumptions entail that it is not the case that you ought to X, even though there is most reason to X. Hence, if these cases are described correctly, and these plausible principles are true, then ANALYSIS OF OUGHT is false.

I reply as follows. Take any alternatives, X and Y, both alleged to be permissible in some situation. We continue to assume that our interlocutor is availing herself of the ideology of weighted reasons and the all things considered ought. There are two kinds of case to consider. Either the nett weights of the reasons in favour of both alternatives will be equal. In that case one ought to perform either one. Such cases clearly present no difficulties for the ANALYSIS OF OUGHT. Or one of these actions will be supported by a greater nett weight of reasons than the other. In all cases of this latter sort, I will suggest, one really ought to do the thing supported by the greater nett weight of reasons. I'll present two considerations in support of the latter claim. Firstly I'll discuss a number of different kinds of cases of putatively permissible

⁸³ This principle is perhaps denied by Scalar Consequentialists, cf. Norcross 'Scalar Act-Utilitarianism.' See also Mill, *Utilitarianism*, chapter 2: "Actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness."

alternatives. In certain cases it is clear that one ought to perform the alternative that realises more value. I'll suggest that there is no relevant difference between these cases and more marginal cases in which one might be tempted to deny that one ought to do the slightly better alternative. Afterwards I'll appeal to an analytic connection between the all things considered ought and accountability. It is appropriate to criticise someone who fails to do the thing supported by the greatest nett weight of reasons. The degree of criticism appropriate is proportional to the difference between the nett weight of the reasons supporting whatever you did and those supporting what you ought to have done.

Where would this leave claims about permissibility? That depends on which kind of permissibility we are talking about (whether practical or moral, for instance), and on what we think about principles relating permissibility and the all things considered ought. It is difficult to maintain that some action is *practically* permissible even if you ought to perform an alternative, indeed, even if you ought not perform the putatively permissible action. However it is an open question whether some action is *morally* permissible even if, all things considered, one ought to perform an alternative. Whatever we say about the former question I think we will be compelled to answer in the affirmative to the question about moral permissibility, for sometimes one ought to do something for prudential reasons *in spite of the fact* that the moral reasons favour an alternative.

Let's start with a familiar case. The current distribution of resources across people in the world is very unequal. Moreover some of this inequality is the result of past and present injustice: imperialism and neoliberal economics. According to most plausible substantive ethical theories, the very richest of the rich (which will almost certainly include anyone reading this) ought to redirect a hefty chunk of their resources to others. *Giving What We Can* invites one to donate 10% of one's pre-tax income to good causes; perhaps giving in accordance with the advice from a meta-charity like *GiveWell*. It might be maintained that this is supererogatory. It is *permissible* to donate 10% of one's pre-tax income to *GiveWell* but it would also be permissible, say, to spend that money eating out with one's friends. Suppose we assume that the 10% figure would leave one with enough resources to eat out occasionally with one's friends – often enough to sustain flourishing friendships with busy people in a busy city. On the other hand, donating 10% of your pre-tax income, for 40 years, at an average salary of \$80,000 would approximately result in the saving of 118 lives, 5,382 years of healthy living, and providing 98,667 years of school attendance.⁸⁴ I'm happy to submit that one has vastly more reason to do this than to 'try the new Italian place,' and that this is indeed what one ought to do, all things considered.⁸⁵

⁸⁴ Figures from www.GivingWhatWeCan.com

⁸⁵ This case suggests that some of our intuitions about supererogation are simply mistaken. It is a tricky ethical question how an individual should balance advantages to herself over advantages to others. It is clear that one should give blood; less clear that one should donate a kidney. I argued in chapter two that it is a mistake to think reasons should be weighted *impartially*, i.e. that they entail, other things equal, that the wellbeing of strangers is as normatively significant for an individual as her own wellbeing or that of her loved ones. This

One could insist, consistent with this result, that it is morally permissible to eat in the Italian place, that it wouldn't be *wrong* to do so, that it is not the case that one morally ought not do so. I find these claims implausible, but they are orthogonal to my point, which concerns (the practical) ought rather than these moral notions.

RIGHTS AND ENTITLEMENTS

Things start to get more interesting when we consider cases in which others have duties or obligations or commitments to do things for us. In such cases it seems that we are permitted to make different choices, even on the assumption that one will be better than others. Here are a few cases:

VOLUNTEER: Suppose Tom offered to pick you up at the airport. You are at your connection in Atlanta, and you remember that Tom has a work deadline in a couple of days, and it would be inconvenient for him to come all the way to JFK to pick you up. But you have heavy bags, and the only way you could get them into town would be to hire an expensive taxi, which you can't really afford.

issue will clearly have implications for supererogation, for it will follow from a more partialist account that sometimes one ought to do something that will make things worse overall, since it makes things better for oneself or one's loved ones. Here we are concerned with broader conceptual issues, so we'll leave this substantive question to the side.

REPARATION: Jerry ate your yoghurt. It wasn't an accident. He was hungry, knew the yoghurt wasn't his, couldn't be bothered to go to the nearby shop, and ate it. Afterwards, he promises to replace it.

FORGIVENESS: You and Bob are colleagues at an advertising agency. Bob was your best friend, but he took advantage of your trust in him, went behind your back, and secured for himself a promotion at your expense. Not long afterwards, in an unrelated matter, he needs you to do him a favour in order to secure some considerable advantage for himself.

In each of these three cases you have a choice. You can choose to relieve Tom of his obligation to pick you up at the airport, you can tell Jerry not to worry about replacing the yoghurt, and you can help Bob. Moreover it is *prima facie* plausible that you are *permitted* to go either way on these matters. You are entitled to let Tom or Jerry off the hook, and free to help Bob.

My responses to these sorts of cases is the same as the response to more general cases of putative permissions. Either the weights of reasons to do either alternative will be equally balanced, in which case one ought to do either. Or else one of these alternatives really is supported by weightier reasons than the other. We can show this by changing the stakes. Increase the importance of Tom's work deadline. Let his livelihood, his reputation, healthcare for his family depend upon it. Let your bags be only mildly inconvenient and the subway rather speedy. Clearly in this case you ought to make the call and release him from the obligation. (Plausibly he would have more reason to fail

to fulfil his obligation anyway.) Let Jerry be starving, the shop an hour's walk away, your yoghurt an unwanted gift. Let Bob's treachery have been extremely mild, and again on this favour rests his livelihood, etc. Again plausibly in these cases you really ought to let Jerry have the yoghurt, and do Bob the favour. We can alter the cases in the other direction to get the opposite result: it would be only mildly inconvenient for Tom to pick you up, and would save you walking seventy miles with your cases (you live in East Quogue). I submit that this result holds quite generally. Hence we have no counterexample to ANALYSIS OF OUGHT. (We can add the usual reminders about questions of moral permissibility, requirement, etc., being strictly orthogonal to this result.)

IMPERFECT DUTIES

We turn now finally to the class of cases that pose a problem for value-based theories of normativity generally, and which perhaps pose the most interesting problem for our analysis: namely, cases of imperfect duties. In ethical theory, imperfect duties tend to be associated with Kant and related deontological approaches. However there is a phenomenon here which any decent ethical theory needs to accommodate. The distinguishing feature of imperfect duties is that they do not confer obligations or requirements in any particular situation.⁸⁶ Instead they obligate one to perform *enough* of a certain kind of action, or to contribute *enough* to some end, over some reasonable period of time. There are loads of phenomena like this. I have an imperfect duty to spend time visiting

⁸⁶ Except the last one: if one has an imperfect duty to see an opera a month and its the last day of the month and you haven't done so, then you have a perfect duty to do so.

my sick grandmother, to look after my health, to stay informed about the development of my friend's children, to make my partner happy. These are all duties, but they are distinguished by a certain kind of optional character. At any given time, one might think, it is open to me whether or not to do something in partial fulfilment of these duties. We get a similar case with the imperfect duty to see Grandma. Suppose I have seen a lot of Grandma recently. In fact, I went to the nursing home to see her this morning. Nevertheless, this afternoon I find myself with some free time. For sure, Grandma would enjoy *another* visit. My chat might not be the best, but let's suppose it is slightly better than reruns of Coronation Street. Still, it doesn't seem to be the case that I *ought* to go back to the nursing home. It would be, as we say, beyond the call of duty to go back a second time. A third case: I have an imperfect duty to develop my understanding and appreciation of great art. Should I go to the art museum this afternoon? But I've been visiting museums all week, and I sat through *Götterdämmerung* last night.

It is easy to see how cases like this could pose trouble for the ANALYSIS OF OUGHT. Take any action which would partially constitute fulfilment of an imperfect duty, but which is such that at that time the agent is not obligated to perform that action. Plausibly in some such case, the action in question would realise more value than any available alternative – for instance than not buying the flowers, or not going back to the nursing home. Still, we have the sense that it is not the case that you ought to buy the flowers, or that I ought to go back to the nursing home.

Imperfect duty cases strike me as the most problematic for our analysis. They also pose a more general explanatory challenge for ethical theories which explain the normative status of particular actions in terms of the value of the consequences (understood broadly enough to include states of affairs instantiated by the action). If we look just at the particular valuable states of affairs that would be realised by an action, it is very hard to explain how it could not be the case that I ought to visit my Grandma, or at least go to the museum, or how it could not be the case that you ought to buy the flowers. In what follows, I want to vindicate our intuitions about imperfect duty cases, while maintaining that they can be accommodated within a value-based framework, and that they do not constitute counterexamples to our analysis of ought.

The central thought in my response to these sorts of putative counterexamples is the following. The problem arises from a conception of the agent's choice situation consisting of too narrow a focus on the immediate consequences of the action. We need to also consider more general features of the choice situation, and how the action under consideration relates to these more general features.

Let's start by introducing the notion of *contributory value*. A state of affairs has contributory value in virtue of being a significant part of a valuable whole.⁸⁷ Consider the state of affairs obtaining just now of the top left half of the Mona

⁸⁷ Cf. Rae Langton, 'Objective and Unconditioned Value.' It is possible to generate a version of this theory of without recourse to claims about value, simply by talking directly about reasons and their weights rather than values and their amounts. This strikes me as less attractive than the approach in the main text.

Lisa's nose having been painted onto the canvas. This has value in virtue of being part of a whole which is valuable, namely, the whole painting. Or consider the state of affairs of Diego Maradona moving into position for the famous "hand of God" goal against England in the World Cup in 1986. His moving into position has contributory value, in virtue of being a part of a whole which is valuable, namely, the beating of England. A third example. There are two of you on a rowboat, one of you is working the left paddle, the other the right. There is a huge waterfall ahead, and you are both frantically paddling away from it. The state of affairs of your moving the left paddle back and forth is part of a valuable whole, namely your both frantically paddling away from the waterfall. (Suppose this is not merely instrumentally valuable: you thrill in the excitement of surviving near-death experiences.)

Spending your evening listening to Gregorian chanting has contributory value, since it is part of a whole which consists in your having exposure to and understanding of a broad range of cultural and aesthetic styles. Spending time whenever you can with your nephew has contributory value, since if you spend enough time with him, over a long enough period of time, you will come to play an important and valuable role in his life, and you can help with his development. Spending time with your nephew is presumably also valuable for its own sake. So a state of affairs can have contributory value as well as final value. Indeed, this will happen often. Listening to the chanting, having a family dinner, going for a run – these states of affairs are presumably valuable for their own sake, but they also have contributory value, since they are parts of a whole consisting in a deeper engagement with your family, sport, or musical

appreciation. Taking your vitamins on a Tuesday has contributory value, because if you take your vitamins regularly enough, you'll be more healthy. Walking up the stairs with all your bags to your sixth floor apartment has contributory value, since doing so regularly is also good for your health. Importantly, as these examples suggest, states of affairs that have contributory value *may* have no other kind of value.

I offer the following theses:

WEIGHT OF CONTRIBUTORY REASONS: The weight of a contributory reason is directly proportional to the contributory value of the state of affairs that constitutes part of the relevant valuable whole.

CONTRIBUTORY VALUE: The contributory value of some state of affairs is a function of the value of the whole and the significance of this part to the whole.

The first principle is a corollary of the thesis presented above: VALUE-WEIGHT CONNECTION implies WEIGHT OF CONTRIBUTORY REASONS. The notion of 'value' employed in this latter definition is understood broadly enough to include non-basic value, instrumental value, and – most important for our purposes – contributory value.

The newcomer here is CONTRIBUTORY VALUE. However this principle is intuitive and familiar. Hold the significance of some part for some whole fixed, and increase the value of the whole, and you'll increase the value of the part.

The same screw has more contributory value when it is holding together the wing of an Airbus A380, than when it is keeping the door of your shed closed. The more interesting part of this principle in the current context is that if you increase the significance of the part, you'll increase its contributory value – and hence by WEIGHT OF CONTRIBUTORY REASONS increase the weight of your reason to bring it about. If the screw is the last one holding the wing on, and the plane is above the middle of the Pacific, it is rather important. If the plane is retired and on display atop the *Intrepid*, less important.

This has significant normative implications, which help us to accommodate the intuitions behind imperfect duties. Suppose that you have been neglecting your avuncular duties, or that you haven't listened to new music for a while, or that you have skipped out on the family dinner to see friends the last few times you were invited. Then the realisable state of affairs consisting in spending time with your nephew, with world music, with the roast turkey, has more contributory value. By WEIGHT OF CONTRIBUTORY REASONS you have weightier reasons to do these things. Crucially these reasons are distinct from the reasons to do these things provided by the final values realised by doing them, i.e. by spending with family, appreciating the music, etc.

We can be more precise. Assume that one needs to perform ten tokens of some act-type in some determinate period of time. (The act-type could be something very straightforward like taking vitamins, getting cardiovascular exercise, not having a glass of wine with dinner, visiting one's sick grandmother. Let's suppose we are talking about taking vitamins.) Let's say the time frame is the

month of January. Let us suppose that you have the chance to perform this activity once a day, and suppose that the feasibility of doing so is the same on any given day. On the first day of the month, you have thirty one occasions in which to do something ten times, so let the contributory value be $10/31$. Suppose one doesn't take vitamins on the 1st day. On the second day, you have only 30 chances to take your vitamins, so the contributory value of doing so is $10/30$. The actual numbers here don't matter; what matters is that the amount is increasing. Suppose you let ten days go by. On the twelfth day, the contributory value of taking the vitamins is $10/20$. Had you taken your vitamins three times by then, the contributory value of taking them that day would be $7/20$. That is to say, the contributory value increases in indirect proportion to the number of available occasions on which one can do the thing, and in indirect proportion to the number of times in which one has done the thing in the relevant period. This is quite intuitive. To change the example, if you have managed to do a lot of exercise this week, or if you have signed up to do a lot in the next few days, you can take today off. If you get to the 22nd day and you haven't taken any vitamins yet, your imperfect duty becomes a perfect duty.

It will be clear that this appeal to contributory reasons therefore promises to explain how it could be, in some cases, that one has more total reason to do something which realises less immediate final value than some alternative. If you take two actions, both of which realise final value and contributory value, and stipulate that you have been neglecting one of these actions, then even if one realises more value at that time, you might have more reason to do the

other. We can oversimplify this with numbers and summation. Going to hear Gregorian chanting: final value 5, contributory value 7. Going out for dinner with best friend: final value 8, contributory value 3. If everything else is equal, you have most reason to hear the chanting.

As this example suggests, contributory reasons also explain circumstances in which one is in good standing, so to speak, i.e. in which one has been doing well with one's medieval art, one's nephew, and one's family. It can seem as though the fact that one has contributed plenty to this particular ongoing project recently can *reduce* the weights of one's reasons to engage in it now. It was this conception of things that presented the problem, since it doesn't seem as though the final value of the state of affairs changes (it is just as good for your nephew, for your family, etc.). Instead, I submit, what happens when we are in good standing is that these *other* reasons are less weighty. Contributory reasons, like dark matter, quietly make up a decent chunk of the total weights of reasons in favour of any given action. When one is in good standing, the significance of the next state of affairs will be less, and the contributory reason less weighty. This can make a significant normative difference.

SECTION FOUR: CRITICISM

We can imagine an interlocutor thinking the following. "I accept that there is sometimes more reason to do one permissible thing than another. But I am still not convinced that I really ought to do that thing. It still seems permissible to

do the other, and I don't seem to be doing anything wrong by doing this other thing. Isn't that just what it is for something to be permissible? I grant you that you can *define* ought in terms of most reason. But what does that achieve? I simply relinquish the *words* "all things considered ought" to you, and continue on regardless, gaily performing whichever among the permissible acts happens to catch my fancy."

To help our interlocutor to recapture her grip on the all things considered ought, we can exploit a plausible principle relating this ought with the appropriateness of *criticism*, namely the thesis that if one fails to do what one ought to do then some degree of criticism is appropriate.⁸⁸ (This notion is not linked with effort, as with the notion of moral praise and blame above.) Take a few simple cases. You are in a hurry. You need to either take the Williamsburg Bridge or the Brooklyn Bridge to get to work. You hear the Williamsburg Bridge is bumper-to-bumper and the Brooklyn is moving quickly. Perfectly appropriate to criticise you if you don't take the Brooklyn without some other good reason. You know that charity A is rated much more highly than charity B on grounds of effectiveness and importance by *GiveWell*, and you fully respect *GiveWell's* analysis. Nothing else speaks in favour of either one. It is perfectly appropriate to criticise you for giving to B rather than A.

⁸⁸ I am borrowing this idea from recent work by Antti Kauppinen (ms). He thinks that the normativity of the what is normative consists in this relation to the appropriateness of criticism. We needn't accept this further claim.

We distinguish this fact about appropriateness from the fact that this or that situated individual (even the agent) ought to criticise the agent. Perhaps someone will shoot you if you criticise her. We'll also ignore questions about the conditions for having the requisite *standing* to criticise.

Now, one is not *irrational* in these cases, nor is it necessarily *wrong* to do these things. These are threshold notions, and reasons are paradigmatically not threshold notions; they are graded. Neither is ‘most reason’ a threshold notion – in the sense that for there to be ‘most reason’ favouring one action over another, it is not the case that that action needs to be supported by a nett weight of reasons weightier than the runner-up by *some specific threshold*. Hence, if ANALYSIS OF OUGHT is plausible, and the thesis relating ought and criticism is plausible, we should not expect the sort of criticism appropriate for failures to do what one ought to be a threshold notion. But as Jonathan Dancy points out, “this does not mean that there is no style of criticism that is applicable in such a case.”⁸⁹ It will often be “silly” to fail to do what – I allege – one ought to do in a situation with more than one putatively permissible alternative.

Again we can motivate this with a series of cases. When one alternative is vastly better than another, and everything else is equal, one would be crazy not to do the better thing. Let charities A and B be respectively at the top and the bottom of *Givewell*’s list, and gradually move them closer. Suppose A is 1st and B is 2nd on the list; suppose that thereby you know that it would be very slightly better to give to A than to B; still you give to B. It wouldn’t be appropriate to criticise you very much – but still appropriate to criticise you a little.⁹⁰ This

⁸⁹ Albeit Dancy may not be too friendly to my general approach. Cf. his ‘Enticing Reasons.’

⁹⁰ At least, in one sense. You are probably still *morally* praiseworthy. Thanks to Adam Lerner for this point.

‘holding accountable’ will fade to nothing on the margins, as we should expect.⁹¹

There are also many different kinds of reasons to do many different kinds of things – as we have seen there are reasons correlated with obligations, duties, and so forth, as well as more straightforward value-based reasons. Correspondingly we should expect that if some kind of criticism were appropriate upon failure to do what one ought to do, which particular kind of criticism is appropriate will vary considerably from one situation to another. In all cases, though, it is plausible that the degree of appropriate criticism is directly proportional to the difference in nett weight between the action supported by most reason and the one you performed.

This point also helps us with worries about similarly weighted reasons and incommensurability. When two actions seem to be supported by a similar nett weight of reasons, it doesn’t seem to matter much which we do. That’s because it doesn’t matter much. You wouldn’t be apt for much criticism if one goes wrong. You shouldn’t care much about going wrong in cases like this. The upshot is that the question about the precision of the weights of reasons is of little *practical* interest. This is all perfectly consistent with our analysis.⁹²

⁹¹ See Antti Kauppinen *op.cit.* for more on mild forms of criticism, dispositions to criticise, and appropriate counterfactuals.

⁹² Finally, what should we say about the two principles, relating permissibility, ought, and ought not? We should say that one of them is false. I don’t much mind which you reject, though I’m inclined to reject the stronger principle according to which if you ought to X, and Y is an action incompatible with X-ing, then you ought not to Y.

CONCLUSION

Where does all of this leave supererogation? As Urmson pointed out, supererogation has to be understood in terms of other notions. In his case, a supererogatory act is permissible, not required, and in some way excellent. But supererogation can be defined in a number of different ways. We discussed many permutations – perhaps tried your patience with them. We considered actions which are good or praiseworthy in virtue of being effortful. We considered different ways in which an action might be permissible. We compared claims of requirement or permissibility which were relative to specific domains, or more general. None of these notions correlated with most reasons, or correlated with ought, in such a way as to be inconsistent with ANALYSIS OF OUGHT. There are always cases and cases. Sometimes the praiseworthy thing, or the morally required thing, or the thing satisfying the imperfect duty is the thing that ought to be done, and sometimes not. Either way you ought to do whatever you have most reason to do.

Works Cited

- Adams, R. 1976. 'Motive Utilitarianism.' *Journal of Philosophy*. Vol, 73, no. 14.
- Audi, P. 2012. 'Grounding: Towards a Theory of the In Virtue Of Relation.' *Journal of Philosophy*, vol. CIX, no. 12.
- Bader, R. (ms). 'Personal, General and Impersonal Good.'
- Baier, K. 1958. *The Moral Point of View*. Ithaca, Cornell University Press.
- Bentham, J. 1789. *Introduction to the Principles and Morals of Legislation*.
- Bergstrom, L. 1966. *The Alternatives and Consequences of Actions*. Stockholm: Almqvist & Wiksell.
- Blackburn, S. 1992. 'Through Thick and Thin.' *Proceedings of the Aristotelian Society*, Supplementary Volume 66, 285-99.
- Brentano, F. 1889/1969. *The Original of Our Knowledge of Right and Wrong*. London: Routledge & Kegan Paul.
- Broome, J. 1991. *Weighing Goods*. London: Wiley.
- Broome, J. 1993. 'Goodness is reducible to betterness: the evil of death is the value of life', in *The Good and the Economical: Ethical Choices in Economics and Management*, edited by Peter Koslowski and Yuichi Shionoya, Springer-Verlag, pp. 70-84.
- Brown, C. 2011. 'Consequentialize This.' *Ethics*, 121 (4):749-771.
- Brunero, J. *Forthcoming*. 'Reasons as Explanations.' *Philosophical Studies*.
- Burra, A. 2010. *Coercion, Deception, Consent: Essays in Moral Explanation*. PhD Dissertation.
- Bykvist, K. 2009. 'No Good Fit: Why the Fitting Attitude Analysis of Value Fails.' *Mind*, vol.118, 469.
- Bykvist, K. 2002. 'Alternative Actions and the Spirit of Consequentialism.' *Philosophical Studies*, 107: 45-68.
- Chang, R. 1997. 'Introduction.' In Chang, R. (ed). *Incommensurability, Incomparability and Practical Reason*. Cambridge: Harvard University Press.
- Clark, Michael J., and Liggins, D. 2012. 'Recent Work on Grounding.' *Analysis*, forthcoming.

- Cohen, G.A. 2004. 'The Truth in Conservatism.'
- Cohen, G.A. 2008. *Rescuing Justice, Rescuing Equality*. Cambridge: Havard University Press.
- Crisp, R. 2000. 'Particularizing Particularism.' In Hooker & Little (eds). *Moral Particularism*. Oxford: OUP.
- Crisp, R. 2005. 'Values, Reasons and the Structure of Justification: How to avoid passing the buck.' *Analysis* 65 (285):80–85.
- Crisp, R. 2008. 'Goodness and Reasons: Accentuating the Negative.' *Mind* 117 (466):257-265.
- Dancy, J. 2000. 'Should we pass the buck?' *Royal Institute of Philosophy Supplement* 47:159-173.
- Dancy, J. 2000. 'Should We Pass The Buck?'
- Dancy, J. 2003. 'Are There Organic Unities?' *Ethics*, 113 (3):629-650.
- Dancy, J. 2004. *Ethics Without Principles*. Oxford: OUP.
- Dancy, J. 2004. 'Enticing Reasons.' In *Themes from the the Moral Philosophy of Joseph Raz*. Oxford: OUP.
- Dreier, J. 1993. 'Structures of Normative Theories.' *The Monist* 76 (1):22-40.
- Dummett, M. 1964. 'Bringing About the Past.' *Philosophical Review*, vol. 73, no. 3, 338-359.
- Eklund, M. 2011. 'What are thick concepts?'. *Canadian Journal of Philosophy*, vol 41, no.1.
- Enoch, D. 2011. *Taking Morality Seriously*. Oxford: OUP.
- Ewing, A.C. 1939. "A Suggested Non-Naturalistic Analysis of Good," *Mind* 48, pp. 1-22.
- Ewing, A.C. 1947. *The Definition of Good*. London: Routledge and Kegan Paul.
- Feinberg, J. 1961. 'Supererogation and Rules.' *Ethics*, vol. 71, no. 4., p.276-288.
- Feldman, F. 1986. *Doing the best we can*. Dortrecht: Reidel.
- Feldman, F. 1993. 'On the consistency of Act and Motive Utilitarianism: A reply to Robert Adams.' *Philosophical Studies*, 70: 201-210.

- Fine, K. 1994. 'Essence and Modality.' In J. Tomberlin (ed.), *Philosophical Perspectives*, 8. Logic and Language. Atascadero: Ridgeview.
- Fine, K. 1995. 'The Logic of Essence.' *Journal of Philosophical Logic* 24, pp. 241–73.
- Fine, K. *forthcoming*. 'A Guide To Ground.' In Correia, F. & Schneider, B. (eds). *Because*. Cambridge: Cambridge University Press.
- Heuer, U. 2004. 'Raz on Values and Reasons.' In Wallace, Pettit, Scheffler & Smith (eds.), *Reason and Value*. Oxford: OUP.
- Heuer, U. (ms). 'The Buck-Passing Account of Value.'
- Holton, R. 2002. 'Principles and Particularisms.' *Proceedings of the Aristotelian Society*, suppl., 76.
- Hooker, B., Mason, E., and Miller, D.E. 2000. *Morality, Rules and Consequences: A Critical Reader*. Edinburgh: Edinburgh University Press.
- Horgan, T. and Timmons, M., 2010. 'Untying a Knot from the Inside Out: Reflections on the 'Paradox' of Supererogation.' *Social Philosophy and Policy*, 27: 29–63.
- Humberstone, L. 1991. 'Two Sorts of Agent-Relativity,' *Philosophical Quarterly* 41: 144–66.
- Hurka, T. 2001. *Virtue, Vice and Value*. Oxford: OUP.
- Hurka, T. 2011. *The Best Things In Life*. Oxford: OUP.
- Hooker, B. & Stratton-Lake, P. 2006. 'Scanlon Versus Moore on Goodness.' In Terry Horgan & Mark Timmons (eds.), *Metaethics After Moore*. Oxford: OUP.
- Jackson, F. 1991. 'Decision-theoretic consequentialism and the nearest and dearest objection.' *Ethics*. 101(3): 461-482.
- Kagan, S. 1988. 'The Additive Fallacy.' *Ethics*, vol.99, no.1.
- Kagan, S. 2000. 'Evaluative Focal Points.' In Hooker et.al.
- Kaupinnen, A. (ms) 'Ought-First.'
- Kelly, T. 2003. 'Epistemic Rationality as Instrumental Rationality: A Critique.' *Philosophy and Phenomenological Research*, 66 (3):612–640.
- Kolodny, N. (ms). 'Instrumental Reasons.'

- Lang, G. 2004. 'A Dilemma for Objective Act-Consequentialism.' *Politics, Philosophy, and Economics*.
- Langton, R. 2007. 'Objective and Unconditioned Value.' *Philosophical Review* 116 (2):157-185 (2007).
- Lechler, A. 2012. 'Do Particularists Have a Coherent Notion of a Reason for Action?' *Ethics*, Vol. 122, No. 4, pp. 763-772
- Little, M. and Lance, M., 2004. 'Defeasibility and the Normative Grasp of Context.' *Erkenntnis* 61 (2-3):435-455.
- Louise, J. 2004. 'Relativity of Value and the Consequentialist Umbrella.' *Philosophical Quarterly* 54 (217):518-536
- Louise, J. 2006. 'Right Motive, Wrong Action: Direct Consequentialism and Evaluative Conflict.' *Ethical Theory and Moral Practice*, 9 (1):65 – 85.
- McDowell, J. 1998. *Mind, Value, and Reality*. Cambridge: Harvard University Press.
- McNaughton, D. and Rawling, P., 1991. 'Agent-Relativity and the Doing-Happening Distinction.' *Philosophical Studies*, 63: 167–185.
- McNaughton, D. & Rawling, P. 2000. 'Unprincipled Ethics.' In Hooker, B. & Little, M. *Moral Particularism*. Oxford: OUP.
- Moore, G.E. 1903. *Principia Ethica*. Online here: <http://fair-use.org/g-e-moore/principia-ethica/>
- Nagel, T. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.
- Norcross, A. 2006. 'Scalar Act-Utilitarianism.' In Henry R. West (ed.) *Blackwell Guide to Mill's Utilitarianism*.
- Olson, J. *forthcoming*. 'Buck-Passing Accounts.' in H. La Follette (ed.), *International Encyclopaedia of Ethics*. Oxford: Blackwell.
- Ord, T. (manuscript) 'How to be a consequentialist about everything.'
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Parfit, D. 2011. *On What Matters*. Oxford: OUP.
- Paz, O. 'Piedra de Sol' or 'Sunstone.' Translated by Eliot Weinberger and Muriel Rukeyser.

- Pettit, P. 1991. 'Consequentialism.' In Singer, P., ed. *A Companion To Ethics*. London: Blackwell Publishing.
- Pettit, P. 1997. 'The Consequentialist Perspective.' in *Three Methods of Ethics*, by M. Baron, P. Pettit, and M. Slote. Oxford: Blackwell.
- Pettit, P. & Smith, M. 2000. 'Global Consequentialism.' In Hooker, B., Mason, E., and Miller, D. E., 2000. *Morality, Rules, and Consequences*, Edinburgh: Edinburgh University Press.
- Portmore, D. 2001. 'Can an Act-Consequentialist Theory Be Agent-Relative?' *American Philosophical Quarterly* 38: 363-377
- Portmore, D. 2003. 'Position-Relative Consequentialism, Agent-Centered Options, and Supererogation.' *Ethics*, vol. 113, No. 2, ps. 303-332.
- Portmore, D. 2008. 'Are Moral Reasons Morally Overriding?' *Ethical Theory and Moral Practice*, 11 (4):369 - 388
- Railton, P. 1984. 'Consequentialism, Alienation, and the Demands of Morality.' *Philosophy and Public Affairs*, vol. 13. , no.2, p.134-171.
- Raz, J. 1999. *Engaging Reason*. Oxford: Clarendon Press.
- Raz, J. 2012. *From Normativity to Responsibility*. Oxford: OUP.
- Raz, J. 2014. 'The Weight of Reasons.' Forthcoming in Lord & Maguire, eds. *Weighing Reasons*. Oxford: OUP.
- Reisner, A. 2009. 'Abandoning the Buck-Passing Analysis of Final Value.' *Ethical Theory and Moral Practice*. 12 (4): 379 – 395.
- Rosen, G. 2010. 'Metaphysical Dependence: Grounding and Reduction.' In Hale, B. (ed.) *Modality*. Oxford: OUP.
- Ross, W.D. 1930. *The Right and the Good*.
- Scanlon, T.M. 1990. 'Promises and Practices.' *Philosophy and Public Affairs* 19 (3):199-226.
- Scanlon, T.M. 1998. *What We Owe To Each Other*. Cambridge: Harvard University Press.
- Scanlon, T.M. 2002a. 'Reasons, Responsibility and Reliance: Replies to Wallace, Dworkin and Deigh.' *Ethics* 112, no.3, 507-528.
- Scanlon, T.M. 2002b. 'Rawls on Justification.' In *The Cambridge Companion to Rawls*. Cambridge: CUP.

- Scanlon, T.M. 2003. *The Difficulty of Tolerance*. Cambridge: Cambridge University Press.
- Schafer, K. (ms). 'Moral Modesty.'
- Schroeder, M. 2007. *Slaves of the Passions*. Oxford: OUP.
- Schroeder, M. 2007. 'Weighting for a Plausible Theory of Reasons.' *Noûs* 41 (1):110–132.
- Schroeder, M. 2011. 'Buck-Passer's Negative Thesis.' *Philosophical Explorations* 12 (3):341–347.
- Schroeder, M. 2007. 'Teleology, Agent-Relative Value, and 'Good.' *Ethics* 117: 265–295.
- Scheffler, S. 1994. *The Rejection of Consequentialism*. Cambridge: Cambridge University Press.
- Sen, A. 1982. 'Rights and Agency.' *Philosophy and Public Affairs* 11 (1):3–39.
- Sidgwick, H. 1874. *Methods of Ethics*. Online here: <http://archive.org/details/methodsethics03sidggoog>
- Singer, P. 1972. "Famine, Affluence, and Morality." *Philosophy and Public Affairs*, 1:229–243.
- Skorupski, J. 2010. *The Domain of Reasons*. Oxford: OUP.
- Smart, J.J.C. 1973. 'An outline of a system of utilitarian ethics.' In *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.
- Smart, J.J.C. and Williams, B. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.
- Smith, M. 2009. 'Two Kinds of Consequentialism.' *Philosophical Issues*, vol.19, issue 1, p.257–272.
- Smith, M. 2003. 'Neutral and Relative Value After Moore.' *Ethics*, vol. 113, No. 3, p.576–598.
- Smith, M. (forthcoming). 'Constitutivism about Reasons.'
- Steglich-Petersen, A. 2011. 'How to be a teleologist about epistemic reasons.' In Steglich-Petersen & Reisner, eds., *Reasons for Belief*. Cambridge: CUP.
- Suikkanen, J. 2005. 'Reasons and Value – In Defence of the Buck-Passing Account.' *Ethical Theory and Moral Practice* 7: 513–535, 2004.

- Suikkanen, J. 2009. 'Buck-Passing Accounts of Value.' *Philosophy Compass*. 4/5: 768-779.
- Thomson, J.J. 1997. 'The Right and the Good.' *Journal of Philosophy*, vol. 94, Issue 6, 273-298.
- Urmson, J.O. 1958. 'Saints and Heroes.' in *Essays in Moral Philosophy*, A. Melden (ed.), Seattle: University of Washington Press.
- Urmson, J.O., 1971, 'A Defence of Intuitionism.' *Proceedings of the Aristotelian Society* 75:111-119.
- Vallentyne, P. 1988. 'Gimmicky Representations of Moral Theories.' *Metaphilosophy* 19 (3-4): 253-263.
- Väyrynen, P. 2006. 'Resisting the Buck-Passing Account of Value.' In Shafer-Landau, R. (ed.), *Oxford Studies in Metaethics, Volume 1*. Oxford: Oxford University Press.
- Väyrynen, P. 2012. 'Thick Concepts: Where's Evaluation?' In Shafer-Landau, R. (ed.), *Oxford Studies in Metaethics, Volume 7*. Oxford: Oxford University Press.
- Vong, G. (manuscript). 'Making Betterness Behave: Constructing a Transitive Deontic Ordering from a Non-transitive Axiological Ordering.'
- Wallace, R.J. 2002. 'Scanlon's Contractualism.' *Ethics*, vol. 112, no.3.
- Wallace, R.J. 2010. 'Reasons, Values and Agent-Relativity.' *Dialectica* 64 (4):503-528.
- Way, J. 2013. 'Value and Reasons to Favour.' *Oxford Studies in Metaethics*, vol. 9.
- Wedgwood, R. 2009, 'Intrinsic Value and Reasons for Action.' In Ernest Sosa & Enrique Villanueva (eds.), *Metaethics*. Wiley Periodicals, Inc. (2009)
- Wedgwood, R. 2009. 'The "Good" and the "Right" Revisited.' *Philosophical Perspectives*. 23, 499-519.
- Wedgwood, R. 2011. 'Defending Double Effect.' *Ratio*, 24. 384-401.
- Wedgwood, R. (manuscript). 'A Platonic Theory of Reasons for Action.'
- Wedgwood, R. (forthcoming). 'The Weight of Moral Reasons.' Forthcoming in *Oxford Studies in Normative Ethics*.
- Wiland, E. 2012. *Reasons*. New York: Continuum Press.

Williams, B. 1973. 'A Critique of Utilitarianism.' In *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.

Williams, B. 1981. *Moral Luck*. Cambridge: Cambridge University Press.

Williams, B. 1985. *Ethics and the Limits of Philosophy*. London: Fontana Press/Collins.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.

Woods, J. & Baker, D. (manuscript). 'Inconsistent Contents and Inconsistent Attitudes: A Defense of B-type Inconsistency.'

Ziff, P. 1960. *Semantic Analysis*. Cornell: Cornell University Press.

Zimmerman, M. 2011. 'Partiality and Intrinsic Value.' *Mind*, vol. 120, 4