Reasons and Normativity

Green Werkmäster, Jakob

2019

[Link to publication](#)

Total number of authors:
1

Reasons and Normativity

JAKOB GREEN WERKMÄSTER
DEPARTMENT OF PHILOSOPHY | LUND UNIVERSITY

Reasons and Normativity

Reasons and Normativity

Jakob Green Werkmäster

LUND
UNIVERSITY

Organization		Document name	
Organization LUND UNIVERSITY		**Document name** DOCTORAL DISSERTATION	
Department of Philosophy Box 190 221 00 LUND Sweden		**Date of issue** 2019-10-26	
Author: Jakob Green Werkmäster		Sponsoring organization	

Title and subtitle : Reasons and Normativity

Abstract

Normative reasons are of constant importance to us as agents trying to navigate through life. For this reason it is natural and vital to ask philosophical questions about reasons and the normative realm. This thesis explores various issues concerning reasons and normativity. The thesis consists of five free-standing papers and an extended introduction. The aim of the extended introdution is not merely to situate the papers within a wider philosophical context but also to provide an overview of some of the central research questions concerning reasons and normativity. The introduction is divided into five chapters. Chapter 1 introduces and outlines the dissertation. Chapter 2 concerns the issue of what a reason is and what different types of reasons there are, such as *pro tanto* reasons and overaall reasons. Chapter 3 discusses the frequently employed weighing metaphor, i.e. that reasons have a certain weight and can be weighed against the weight of other reasons. Chapter 4 covers how reasons relate to other normative notions such as ought, value, and obligation. For example, it discusses whether it is the case that for something to be valuable just is for there to be reasons to favor it. Chapter 5 concludes the introduction by briefly exploring how it all relates.

The thesis can be said to be embedded in a recent research trend within philosophy of normativity that has a distinguished focus on reasons.

The five papers deal with various issues concerning reasons and normativity. Paper I argues for a novel and theoretically parimonous way to understand background conditions for values and reasons. Paper II explores an interpretation of the distinction between subjectivism and objectivism about resons and value so as to accommodate that the distinction does not committ either party to certain first order claims about what reasons there are or what makes objects valuable. Paper III discusses the dictum 'ought implies can' and how to properly interpret the 'can'. Paper IV defends principles according to which what we ought and have reasons to do transmit from ends to the necessary means of that action. Paper V analyzes how the Fitting Attitudes analysis of value should best understand degrees of value.

Key words
Reasons, Normativity, Value, Ought, Buck-Passing

Classification system and/or index terms (if any)

Supplementary bibliographical information	**Language** English
ISSN and key title	**ISBN** 978-91-88899-59-0 (print)

Recipient's notes	**Number of pages 230**	Price
	Security classification	

Signature   Date 2019-08-27

Reasons and Normativity

Jakob Green Werkmäster

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into the context and summarizes the main points of the papers. Then, the research publications are reproduced. The research papers may either have already been published or are manuscripts at various stages (in print, submitted, or in draft).

To Barbro Werkmäster

Table of Contents

List of Papers

Paper I: **Towards a Reasons-First View of Background Conditions**

Andrés G Garcia (co-author) and **Jakob Green Werkmäster** (co-author)

Submitted manuscript

Paper II: **Subjectivism and the Framework of Constitutive Grounds**

Andrés G Garcia (co-author) and **Jakob Green Werkmäster** (co-author)

Ethical Theory and Moral Practice (2018) 21:155. doi:10.1007/s10677-018-9862-1

Paper III: **A Fair Reading of 'Ought Implies Can'**

Jakob Green Werkmäster

Submitted manuscript

Paper IV: **Normative Transmission and Necessary Means**

Jakob Green Werkmäster

Philosophia (2019) 47:555. doi: 10.1007/s11406-018-9976-7

Paper V: **How Valuable is It?**

Henrik Andersson (co-author) and **Jakob Green Werkmäster** (co-author)

Submitted manuscript

Paper II and IV are produced with the kind permission of Springer.

Acknowledgements

Dear reader,

I am so happy that you get to read my acknowledgments in my thesis. There are so many people that I wish to thank. It is regrettable that there is a prevailing myth within philosophy according to which philosophy gets written in solitude and isolation. Without the people around me there would be no thesis.

I probably owe more than I realize to my two supervisors, David Alm and Toni Rønnow-Rasmussen. Throughout this process I've felt that your doors have been open whenever I needed you. It made me feel safe. Thank you. I am grateful to David for his fascinating mix of blunt cynicism and unwavering support. Thank you Toni. Thank you for introducing me to the philosophy of reasons and normativity when I was a bachelor student. Thank you for all the words of encouragement and discussions (philosophical and otherwise) throughout the years.

Henrik Andersson and Andrés Garcia, few things in life has been as deeply satisfying and philosophically stimulating as having the opportunity to write papers with you. Thank for your friendship. Thank you for showing me just how fun (and painful) writing philosophy can be. Thank you for sharing an office with me for three years.

Jonas Olson gave me invaluable comments on the penultimate draft of this thesis when he was my opponent at the final seminar. His comments have been vital in developing this thesis, everything from killing my darlings to killing my darlings.

I admire Marta Johansson and Mattias Gunnemyr for their thorough readings of later drafts of the thesis. Their feedback and discussions added much clarity to my writing.

I am also thankful for the input and discussions throughout the years with my PhD colleagues: Henrik Andersson, Eric Brandstedt, Andrés Garcia, Cathrine Felix, Fritz-Anton Fritzson, Marta Johansson, Mattias Gunnemyr, Frits Gåvertsson, Seyyed Mohsen Eslami Ardakani, Gloria Mähringer, Jeroen Smid, Signe Savén, Ylwa Sjölin Wirling.

The breeding ground for many early drafts and ideas presented in this thesis has been ventilated at the higher seminar in practical philosophy in Lund. I am

Werkmäster, thank you for never thinking that a career in philosophy made you cool.

Marta Johansson, I need you to come now. I need you to come now straight away. Bring the calculator. Bring the Steinway Piano. I love you. Later.

Finally, I want to thank my grandparents Jerk-Olof Werkmäster and Barbro Werkmäster. Thank you for your sincere curiosity and for giving a small child the extraordinary experience of having everything they say being taken seriously.

<div align="right">Jakob Green Werkmäster.</div>

Introduction

Some facts seem to stand in a special relation to us in that they call for a response. Some facts are normative reasons.[1] For instance, the fact that I made a promise to buy groceries is a reason for me to do it. Regardless of what situation we find ourselves in, we have reasons to perform certain actions, hold certain beliefs, and have certain attitudes (e.g. feelings, desires, wants, and other non-doxastic attitudes). In short, we have reasons *for* a lot of things. Reasons can be more or less strong or weighty in the sense that, for example, some speak heavily in favor of an action while others only give marginal support for that action. More often than not, our reasons concern mundane situations such as deciding where to go for lunch. Even in these trivial choice situations we are dealing with reasons for and against going for lunch at one place rather than another.[2] We can suppose that the fact that one restaurant is cheap is a reason to have lunch there, and the fact that the other restaurant is cozy is a reason to have lunch there. At other times, however, our choices are of a more serious nature. Consider, for example, the question as whether to raise taxes or reduce funding for the arts and humanities.

Normative reasons have some sort of sway over us in the sense that reasons matter. In fact, it may be that reasons are what make matters matter. Normative reasons are, in a wide sense, action-guiding and in some way connected with what we ought to do.[3] It is of both theoretical and practical importance that we understand

---

[1] Henceforth I will refer to 'normative reasons' as 'reasons' unless the distinction between normative and other sorts of reasons is needed.

[2] Philosophy has a habit of focusing on high-stake cases where it is of the outmost importance that we make the right choice, such as with climate change or lying to murderers. However, this habit should not lead us to deny that reasons are part and parcel of our daily lives.

[3] I here follow Andrew Reisner (2015b) and use 'do' in a slightly idiosyncratic manner as a universal verb, or a placeholder, meant to include performing certain actions, having certain attitudes, or having certain beliefs. I will throughout the thesis be explicit when I intend to capture only actions by 'do'.

how reasons function, and how what we have reasons to do connects with what we ought to do and what is valuable.

In this dissertation the focus lies on the structure of reasons and normativity. The focus is neither on classic meta-ethics nor on substantive normative ethics. It is somewhere in between, on the structural level. The goal of the thesis is to deepen our understanding of normativity in general and normative reasons in particular. The thesis consists of two parts: An extended introduction (divided into four chapters) and five papers. The extended introduction aims to give a rough outline of the structure of reasons and normativity, touch upon some of the important debates concerning reasons and normativity, and situate the papers within these debates. The papers are self-contained, but they are intended as journal articles and thus require some background knowledge and familiarity with the field. The extended introduction aims to provide some of that background knowledge. In addition, the extended introduction can be seen as something that has so far been lacking in the literature: An overview of some of the central concepts and properties that play a prominent role in the literature on the nature of reasons and normativity.[4]

Before we begin proper, I want to introduce some of the more rudimentary distinctions that need to be made in a discussion of reasons. The explication of these distinctions occupies most of the extended introduction, but given their sheer number I think it is useful to situate them in a rough taxonomy at the outset.

We get at the first distinction by answering the question: "What is x a reason for?" We can have reasons either for holding certain beliefs, doing certain actions, or having certain attitudes.[5] I call these three different reasons theoretical reasons, practical reasons, and evaluative reasons. For instance,

---

[4] The papers are journal articles, and I apologize for occasional repetitions – e.g. the plentiful introductions to the Fitting Attitudes Analysis of value (FA analysis). The FA analysis (which we will look at more carefully in the chapter "On Buck-Passing") analyzes an object being valuable in terms of it being fitting to favor the object.

[5] Instead of the phrase "reasons for attitudes" other philosophers use phrases such as "reasons for feelings" (Skorupski, 2010a) or "reasons for pro- or con-attitudes" (Rabinowicz & Rønnow-Rasmussen, 2004). I prefer the phrase "reasons for attitudes" because it is more neutral. By it I mean reasons for feelings, desires, emotions, wants, and other non-doxastic attitudes. One of the big drawbacks of using "reasons for attitudes" is that beliefs are also a sort of attitude, and I do not include beliefs (or suspensions of belief or disbelief) under the phrase. I believe that this

Theoretical reasons: My visual experience of a cup on my desk is (in most cases) a reason for me to believe that there is a cup on my desk.[6]

Practical reasons: My promise to buy groceries is a reason for me to buy groceries.[7]

Evaluative reason: My failure to keep my promise to buy groceries is a reason for my friend to be annoyed with me.[8]

The second distinction I want to make emerges when we answer the question: "What kind of reason is it?" For instance, it could be a moral, an aesthetic, or a prudential reason. The three-way categorization of moral, aesthetic and prudential reasons should not be read as an exhaustive list, nor do I intend to take a stance on what kinds of reasons there actually are. My goal is, more accurately, that of making sense of what it is to be a reason of a certain kind. To illustrate the distinction, some might say that my promise to my friend to read one of her papers is a moral reason to do it. They might add that the fact that I could work on my own paper instead of reading her paper is a prudential (and not moral) reason *not* to not read her paper; and that my friend's eloquent writing style is an aesthetic reason to read her paper. What the modifications 'moral', 'prudential' and 'aesthetic' signal is what *kind* of reason is in play. They say something about the normative source of the relevant reason. Exactly what is meant by *source* is something I will go into in greater detail in the final chapter of the introduction, "How It All Relates". For now I think it suffices to say that the source of a reason is an explanation of *why* a particular fact has the property of being a reason.

---

irregularity in the usage of "attitude" is compensated for by the fact that it is both short and simple, and also theoretically more neutral than other phrases. Depending on one's view on feelings and emotions, not everything that would fall within this third category of reasons can properly be described as a feeling.

[6] Theoretical reasons need not be limited to reasons for beliefs; they may include reasons to revise one's credence in a certain belief, or depending on how you understand beliefs, reasons to abstain from forming a belief at all. Thanks to Jonas Olson for pointing this out.

[7] Throughout the thesis I will assume that abstaining from doing a certain action is also a form of action. Regardless of whether one imposes a hard distinction between action and omissions, we will have practical reasons to omit doing things.

[8] While this thesis' main focus is on reasons for actions, and to some extent reasons for attitudes, many aspects of the discussion are directly translatable to discussions of theoretical reasons

The third distinction is the one between *pro tanto* reasons and overall reasons.[9] We ordinarily talk about particular reasons for and against, say, a certain action, but we also talk about what we have reason to do overall – what the case is for performing one action rather than another. The particular reasons are *pro tanto* reasons. *Pro tanto* reasons function at the contributory level; they speak in favor of, say, an action in a certain aspect to a certain extent. Overall reasons, as the name suggests, function at the aggregated level. For instance, the fact that buying groceries is boring is a *pro tanto* reason against doing it, but the fact that I promised to do so is a reason to buy groceries, as is the fact that I want to use the groceries to make dinner. At the overall level, my reason for buying groceries is stronger than my reason against doing so.

The first chapter of the extended introductory chapter goes into more detail on the distinction between *pro tanto* reasons and overall reasons.

I am interested in normativity. Exactly how to describe what normativity is, or what is so normative about 'normative' reasons is notoriously difficult. That said, I think it will be helpful to begin by contrasting normative reasons with other varieties of reason – if we cannot say what normative reasons are, we can at least make it a little clearer what they are not. I distinguish between three varieties of reasons:

> *Normative reasons*: A normative reason is that which counts in favor of something, such as a promise to do something.
>
> *Motivating reasons*: A motivating reason is the state of affairs, or propositional content of the agent's attitudes and beliefs, (i.e. *what* the agent desires or believes), which motivates the agent to act.
>
> *Explanatory reasons*: An explanatory reason is whatever explains why the agent acted, or more generally what explains why something is the case (e.g. the agent's beliefs, a society's norms and culture, and so forth).

An example, borrowed from Toni Rønnow-Rasmussen, helps to illuminate the distinctions:

---

[9] Instead of 'overall' some philosophers prefer the term '*all-things-considered*'. For more on *all-things-considered* see footnote 32 p. 32 and footnote 38 p. 39.

An agent *a* is asked why she jumped into the water and says something like 'The reason is that God told me to save the person from drowning' (Rønnow-Rasmussen, 2011, pp. 133 - 134).

In this case the explanatory reason, the thing that explains why she jumped into the water, is *a*'s belief that God told her to do so (and perhaps her desire to do what she believes God told her to do). Her motivating reason on the other hand, is not the belief that God told her to do so, but rather the content of that belief: [God told her to do so]. Sometimes one and the same thing is a normative reason *and* that which motivates the agent, but this is not necessarily so.

So, the concept of a normative reason is not to be confused with the concept of a motivating reason or the concept of an explanatory reason. But what is a normative reason? What is the normative sense of the word 'normative'? Unfortunately, I do not have a positive characterization of this other than to say that the normative sense of 'normative' that I will be using means 'having to do with (normative) reasons or (normative) oughts'. Such a definition of normativity is of course unhelpful and circular unless we already have a prior understanding of normative reasons or normative oughts. Fortunately, I believe that most readers are already competent to distinguish between the truly normative and other senses of 'normative', 'reasons', and 'oughts'.

To make matters worse, not all reasons or oughts that appear to be normative are 'truly' normative.[10] Moral reasons and moral oughts are paradigmatically normative, but etiquette-based reasons or oughts are not 'truly' normative. John Broome formulates the problem as follows: "Even the word 'normative' has a non-normative (in my sense) sense" (Broome, 2013, p. 11). To exemplify this he goes on to say that the word 'normative' "may be used to mean 'to do with norms', where a 'norm' refers to an established practice or alternatively to a rule or requirement" (Ibid). The fact that according to etiquette the dessert spoon should be placed above the plate is a reason to do so, but it does not 'truly' count in favor

---

[10] 'Ought', being a modal auxiliary verb, creates friction between philosophical clarity and the English language. Grammatically speaking, it is incorrect to write 'oughts' as the plural of ought or to treat 'ought' as a noun in general. However, grammar should not stand in the way of philosophical clarity. What I refer to by 'oughts' or 'the ought' is the fact that an agent(s) stands in the ought-relation to a certain action, belief, or attitude. For it to be the case that an agent ought to keep a promise is for the agent to stand in the ought-relation to the action of keeping the promise. For more on ought and the English language, see (Chrisman, 2012)

of doing so on its own. One way, perhaps, to approximate 'true' normativity is by reflecting on the following two statements:

> According to etiquette I have a reason to φ, but do I have a reason to φ?

> According to morality I have a reason to φ, but do I have a reason to φ?

The first question has an 'open' feel to it that the second statement does not have – open, in the sense that it is intelligible to answer "no" to the first question but not to the second.[11]

Over the last twenty odd years, normativity, and the concept of a normative reason, have received considerable philosophical attention (Dancy, 1993; Korsgaard, 1996; Broome, 1997; Parfit, 1997; Dancy, 2004; Darwall, 2006; Schroeder, 2007; Parfit, 2011a; Raz, 2011; Broome, 2013; Olson, 2014; Fogal, 2016).[12] It is perhaps not surprising that, in its wake, the surge in research interest and rapid progress has left a research field that is somewhat in disarray, with philosophers using a range of contrasting concepts and terms. Numerous terms now denote the same concept, and few established and rigorous characterizations of the concepts are to hand.[13] For this reason it is quite hard to keep track of all of the sophisticated distinctions that have been made. In this thesis, as far as

---

[11] This can be compared with the way Scanlon (1998, p. 97) and perhaps especially Philip Stratton-Lake (2002, pp. xx - xxi) interpret Moore's (1993/1903) open question argument.

[12] One might say that the renewed research interest in reasons as we know it today really comes into existence with Scanlon (1998). Other potential starting points would be Henry Sidgwick (1981/1874), or perhaps it starts with W. D. Ross (2002/1930) where we first get the prototype formulations of '*pro tanto* reasons' in terms of '*prima facie* duties'. Another potential starting point is Baier (1958), the first to distinguish between explanatory, normative and motivating reasons; or Toulmin (1958), Williams (1981a) or Nagel (1970, 1986), where the first systematic discussions of the nature of reasons appear. Another possible point of inception is Parfit (1984), whose influence shaped more than a generation of philosophical discourse.

[13] I have a list where I record different types of reasons, and the different terms used for these types of reasons, in the work of people I have read. So far, it consists of 91 terms and at least 65 different types of reason, and that is only counting those used by authors with a specific usage in mind. This entire dissertation could consist merely of attempts to make sense of the list and its taxonomy.

possible, I have tried to resist the urge to add to the lexicon of terms being used to denote the same concepts.[14]

There is one more thing that I would like to mention before I outline the remaining chapters of the extended introduction. The thesis contains a lot of examples. These examples should not be read as endorsements of substantive normative claims. Although the thesis has reasons and normativity as its main focus, it does not contain any kind of finger-pointing – it does not prescribe what reasons we actually have or say how we actually ought to live. As far as possible, the thesis remains neutral when it comes to substantive normative issues and does not engage in normative theorizing of the sort one engages in when figuring out whether, for instance, Kantianism or Rossian Pluralism is the correct moral theory.[15] Nothing I say hinges on which specific normative theory turns out to be correct. I plead to readers to read the examples used in the thesis as just that – examples, or better still examples of examples. My interest lies purely with the structural features that the examples exhibit.

Let me now outline the chapters that make up the extended introduction.

These chapters give an overview of, and expand on, the ways in which contemporary philosophers have answered the question "How are we to understand the concept of a normative reason?". They do so by introducing the concept of a reason, examining its formal features and asking what different types of reasons there are. I start with the concept of a reason in my attempt to unpack the structure of normativity because I view reasons as, metaphysically speaking, the most fundamental normative feature on which other normative features are grounded. Additionally – and something that should not be underestimated – a pedagogical case can be made for starting in this way. Reasons are something that we are all more or less familiar with. Anyone who has ever deliberated over options is very likely to have some familiarity with reasons and how they work.

The chapter "The Weighing Metaphor" aims to clarify what it means to say that a reason has a weight, or strength, and how we should understand the metaphor of 'weighing reasons' against one another. It does so by analyzing the weighing metaphor as it is commonly employed in the debate concerning reasons and

---

[14] My own preferred nomenclature would, of course, have been far superior.

[15] It is, of course, some sort of Rossian Pluralism.

normativity. The weighing metaphor says, simply put, that we talk about weighing reasons in the same way we talk about weighing physical objects. A clear example of this usage, drawn from outside philosophical debate, is that of the Lady Justitia scales of justice as they appear in the judicial context, reflecting the weighing of different interests and evidence.

The chapter "On Buck-passing" utilizes some of the conceptual tools developed in the two previous chapters in order to provide a plausible story about how reasons relate to other significant normative features such as value, ought and obligation. Specifically, I argue that we can understand both *being valuable* and *ought* in terms of reasons. The chapter also highlights some commitments and potential problems that are entailed by the understandings that have been recommended, and asks whether we should modify our accounts of value, ought and obligations in light of these.

Following this, the chapter "How It All Relates" concludes the extended introduction by exploring a range of answers to questions about the way normativity is structured by looking at different possible theories of the architecture of normativity. A theory of normativity's architecture is essentially a set of claims indicating what normative domains and sources there are, and how they interrelate.

The Reason-Relation

The goal of this chapter is to characterize the property of being a reason. As we shall see, there are a few different types of reason-relation, and several ways of capturing them. I provide these characterizations first and foremost because in order to understand how reasons work we must better understand what they are.

I take reasons to be facts.[16] Facts are here to be understood as (true) proposition-like entities picked out by 'that-clauses'. For instance, that I promised my friend that I would return a book tonight is a *reason* for me to return the book tonight. That I promised my friend that I would return a book tonight is also a *fact*.[17] I will sometimes use the term 'reason-fact' to refer to facts with the property of being reasons. Reason-facts are no more ontologically queer than any other mundane facts. The fact that I promised to buy groceries is not queerer ontologically than the fact that there is a chair in my office.[18] All reasons are essentially normative in that they are endowed with a certain normative property. The property that all our reasons share is that of *being a reason*. To have the property of being a reason is to stand in a certain relation to, at least, an agent and

---

[16] Other suggestions in the literature include taking reasons to be properties (Brown, 2014). Schroeder (2007) views some reasons as sets and other reasons as facts. Which ontological category reasons belong to makes little or no difference to the arguments in this thesis. The main reason why I choose, in this thesis, to view reasons as facts, and only as facts, is that talking about facts is easier than talking about properties. However, see page 36-37 for some complications.

[17] For some slight elaboration on my view of facts see pp.36-37.

[18] Some believe that facts themselves are ontologically questionable (Betti, 2015). While I disagree with philosophers who banish facts from their ontologies, the important thing to note is that those who find facts problematic do not believe that reasons *qua* facts are any more problematic than other facts.

a response. While facts are arguably ontologically innocent, according to most theorists of reasons, the normative property of being a reason might not be.[19]

At a minimum, the reason-relation exhibits the following structure:

The Minimal Account: R(x, S, φ)[20]

On the minimal account the reason-relation is a three-place relation, the fact (or collection of facts), x, is a reason, R, for an agent (or agents), S, to φ. Here φ is a place-holder for a belief, an action or an attitude.[21] Depending on whether φ takes the form of a belief, an action, or an attitude, we are dealing with a theoretical reason, a practical reason, or an evaluative reason.[22] In this thesis the main focus is on practical and evaluative reasons.[23]

Once we depart from this minimal account, divergent views will start to emerge of the appropriate way to characterize reason-relations.

The tripartite division of reasons into reasons for beliefs (theoretic reasons), reasons for actions (practical reasons) and reasons for attitudes (evaluative reasons) is not universally accepted.[24] Some, for example, (Parfit, 2011a), make do with practical and theoretical reasons alone, and refer to both reasons for actions and reasons for attitudes as practical reasons; others deny the existence of reasons for desire-like attitudes (Maguire, 2017). Like Skorupski, I take the difference

---

[19] For more on the alleged ontological innocence of reasons, see (Shafer-Landau, 2003; Parfit, 2011b; Olson, 2014, 2018).

[20] For convenience I have leave the quantifier implicit, so 'R(x, S, φ)' = There is a fact x such that x is a reason for S to φ = $\exists(x)$ R(x, S, φ).

[21] Remember that by the, perhaps unfortunate, phrase "reasons for attitudes" I refer to certain attitudes, feelings, desires, wants, and other non-doxastic attitudes.

[22] C.f. (Skorupski, 2010a, pp. 36 - 37).

[23] While the minimal account contains the characteristics everyone agrees that the reason-relation has, 'minimal' may not be the optimal name for this relation. Another, perhaps more fitting, name would be 'generic'.

[24] For now, I wish to remain silent on how types of reason relate to kinds of reason – e.g. on whether we can have theoretical moral reasons, practical evidential reasons, and evaluative prudential reasons. In principle, there is conceptual space to allow every type of reason to be every kind of reason. For more on this, see the chapter "How It all Relates".

between these three kinds of reason to be "as clear as the difference between action, belief, and feeling" (Skorupski, 2010a, p. 36).[25]

Pro Tanto Reasons

Usually, in everyday life, when people talk about reasons what they have in mind are probably *pro tanto* reasons. A *pro tanto* reason is, roughly, a fact that speaks in favor of an option with a certain degree of strength. To illustrate: The fact that Alexander Hamilton promised Aaron Burr to water Burr's plants is a *pro tanto* reason for Hamilton to do so. The fact that it would be a mild annoyance for Hamilton to water Burr's plants is a *pro tanto* reason for Hamilton to not do so. Furthermore, we seem to be able to weigh these two *pro tanto* reasons against each other and say which is the stronger, or weightier, reason.[26] The two *pro tanto* reasons certainly seem to differ in strength and plausibly the promise-reason outweighs the mild-annoyance-reason. The *pro tanto* reason relation has roughly the following structure:

> *Roughly, the Pro Tanto Reason Relation*: R(x, S, w, φ)

The fact x is a reason for an agent S with a degree of strength, w, to φ. Another way to express the same idea would be to say that the fact x stands in the reason-relation to an agent, a degree of strength and an action, belief, or attitude.

In order to get a further grip on, and clarify our understanding of, *pro tanto* reasons it will be beneficial to see how they have been characterized by previous commentators. I have chosen to take a closer look at the following four philosophers because they elucidate different aspects of the *pro tanto* reason-relation, because they represent influential stances in the debate, and because their accounts are not straightforwardly compatible (and their potential incompatibility is philosophically interesting). The four philosophers are Jonathan Dancy, John Broome, John Skorupski and T. M. Scanlon.

---

[25] However, in the chapter "How It All Relates" we will have the opportunity to mention other ways to cut the normative cake.

[26] I use "strength of a reason" and "weight of a reason" interchangeably.

In *Ethics Without Principles* (2004) Dancy gives a characterization of pro tanto reasons. Dancy does not actually use the term "*pro tanto* reasons" but instead talks about "contributory reasons". Although terminological diversity is symptomatic of the debate, it should not trouble us that Dancy is using a different term. In a footnote he acknowledges that the way he uses "contributory reasons" is identical to the way others use "*pro tanto* reasons" (Dancy, 2004, p. 17).[27] Dancy's characterization is as follows.

> A contributory reason for action is a feature whose presence makes something of a case for acting, but in such a way that the overall case for doing that action can be improved or strengthened by the addition of a second feature playing a similar role. [...] Contributory reasons are officially reasons capable of doing what they do either alone or in combination with others. But they can combine in peculiar and irregular ways (Dancy, 2004, p. 15).

The quote is interesting, and we can learn a lot from it about how Dancy believes *pro tanto* reasons work. In the passage, Dancy restricts himself to reasons for actions and claims that a *pro tanto* reason "makes something of a case for acting". It is a little unclear what Dancy means by "a case for acting" and an "overall case for doing that action". When we are finished characterizing *pro tanto* reasons we will have the opportunity to revisit Dancy's formulation of an "overall case for acting", but for now let us postpone that discussion.

The crucial thing to note about Dancy's characterization of *pro tanto* reasons here is that while they make a case for doing something, that case can be strengthened or weakened by the presence of other *pro tanto* reasons. We also learn that Dancy believes that (the weight of) *pro tanto* reasons need not combine in a straightforward manner. A *pro tanto* reason supporting a certain action might make a case for performing that action, but the addition of an additional *pro tanto* reason supporting the very same action might make the overall case for performing the action weaker than it originally was. Moving on from Dancy's characterization, but lingering on the feature of *pro tanto* reasons of making a better or worse case for acting, Broome (2013) gives the following definition of *pro tanto* reasons using *N* is a placeholder for an agent and *F* is a universal verb.

---

[27] Another philosopher who uses a different term for *pro tanto* reasons is Skorupski (2010a, p. 33). He uses the term "specific reasons". As far as I can tell, there is no philosophical difference.

> A *pro tanto* reason for N to F is something that plays the for-F role in a weighing explanation of why N ought to F, or in a weighing explanation of why N ought not to F, or in a weighing explanation of why it is not the case that N ought to F and not the case that N ought not to F (Broome, 2013, p. 53).

Broome's definition makes use of the weighing metaphor – *pro tanto* reasons are essentially things we weigh. To make sense of the definition we need to briefly clarify what Broome means by a 'weighing explanation'. First, take a standard case of weighing two objects where one object outweighs the other. According to Broome, a normal weighing explanation of why one object outweighs another goes something like the following:

> There is at least one object in the left-hand pan, and there may be one or more objects in the right-hand pan. Each object has a weight. The combined weights of the object in the left-hand pan exceed the combined weights of the objects in the right-hand pan. That is why the balance tips left (Broome, 2013, p. 52).

Broome suggests something analogous happens in the case of reasons. We put the *pro tanto* reasons in favor of F in one pan, and the *pro tanto* reasons against F in the other pan. If the combined weight of the *pro tanto* reasons for F is greater than the combined weight of the *pro tanto* reasons against F, then we have a weighing explanation of why we ought to F. In Broome's definition we are given an answer to the question how 'reasons' and 'oughts' are related – *pro tanto* reasons are what figure in the for-F-role and the against-F-role in a weighing explanation of why it is the case that we ought to do F (or not do F).

I agree with Broome in so far as I also believe that *pro tanto* reasons play a role in explaining what we ought to do. I am skeptical, however, about Broome's characterization of a *pro tanto* reason. Can we understand the "for-F role" independently of understanding what a reason is, or what it is for a fact to "count in favor of responses"? Like other philosophers (Skorupski, 2010a, p. 50 ff; Brunero, 2013; Dancy, 2015, p. 181; Kearns & Star, 2015, pp. 237 - 238), I am not convinced this is possible. Broome argues that the "for-F role" simply is the concept of weight, and that we have a prior understanding of weight. Kearns & Star (2015, p. 238) point out that while we might have a prior understanding of weight that is independent of "counting in favor" when it comes to non-normative mechanical weighing, it is less clear that we have an independent understanding of *normative weight*.

A reason (or collection of reasons) can either be conclusive, sufficient and insufficient.

> *Conclusive Reason*: An agent has a conclusive reason to φ if and only if the reason to φ is stronger than the reasons (taken together) to not-φ.

> *Sufficient Reason:* An agent has sufficient reason to φ if and only if the reason to φ is at least as strong as the reasons (taken together) to not-φ.

> *Insufficient Reason*: An agent has insufficient reason to φ if and only if the reason to φ is weaker than the reasons (taken together) to not-φ.

Unlike *pro tanto* reasons these reason relations are strict and non-weighted. In saying this, I mean that they do not come in degrees of strength. There are no very conclusive reasons, or weak sufficient reasons.[28]

While these definitions might look obvious and intuitive, there are at least two philosophers who, for different reasons, disagree with the above characterization. Scanlon (2014) disagrees on the grounds that my characterization is backwards – according to him, we should understand the strength of reasons in terms of sufficient reason. Skorupski (2010a), on the other hand, would dispute the characterization because he denies that sufficient reasons and conclusive reasons are definable in terms of weighted non-strict notions of reasons.

---

[28] So even if, for instance, the sufficient reason relation does not come in degrees, of course the underlying reason does have a degree of strength. It is like when Usain Bolt wins the 100m race. The property of winning the race does not come in degrees, but of course he ran the race with a certain degree of speed in virtue of which he won it. Moreover, if Bolt won race$_1$ with a certain time and won race$_2$ with a slower time, there would be a sense in which it is true that the first time is better than the latter. Bolt in race$_1$ would have won against Bolt in race$_2$. However, this does not mean that he won the first race *more* than he won the second race. He won both races equally. Likewise, where a conclusive reason would not have been conclusive in different circumstances, this does not mean that the reason is conclusive to a higher or lower degree.

To show how my position differs from theirs, I shall look more closely at the ways in which Skorupski and Scanlon characterize the *pro tanto* reason relation.

> *Skorupski's Account*: R(x, t, S, w, φ) (2010a, p. 37).

> *Scanlon's Account*: R(x, S, c, φ) (2014, p. 37).

So, for Skorupski the reason-relation is a five-place relation: The fact (or collection of facts), x, is at time, t, a reason of degree of strength, w, for the agent, S, to φ. For Scanlon, it is a four-place relation: The fact, x, is a reason for, S, in circumstances, c, to φ.

The main difference between the two is that while Skorupski's account has the fine-grained relata of degree of strength and time, Scanlon opts for the thicker relatum of circumstance. So, for Skorupski all *pro tanto* reasons are reasons at specific times with certain degrees of strength to do things. For Scanlon all reasons are reasons only in specific circumstances.

On Scanlon's account it only makes sense to talk about the strength of reasons in relation to other reasons. We can understand a *pro tanto* reason being stronger than another if it is a sufficient reason (SR) but the other one is not. In other words, a reason is stronger than another if the following holds:

> One consideration, x, outweighs another, y, if the following hold: R(x, S, c, φ), R(y, S, c, ψ), where ψ is a course of action incompatible with φ, and SR(x, S, c, φ) but not SR(y, S, c, ψ), although SR(y, S, c', ψ) where c' is a set of circumstances as normatively similar to c as possible except that x does not obtain in c' (Scanlon, 2014, p. 108).[29]

It can be seen, then, that Scanlon treats the relation of being a *sufficient reason* as more fundamental than weighted notions of reasons. A *pro tanto* reason x is stronger than a *pro tanto* reason y if and only if, in c, x is a sufficient reason and y is not a sufficient reason, although y is a sufficient reason in circumstances c' that are similar to c with the difference being that x does not obtain in c'. On this view, it does not make sense to talk about x being a strong or weak *pro tanto* reason; it

---

[29] Note that for the purpose of readability I have exchanged Scanlon's original annotation of R(p, x, c, a) for the one I have used in the text in order to avoid using the same symbol for different variables.

only makes sense to talk about whether x is stronger or weaker than another *pro tanto* reason. It seems to be an implication of Scanlon's view that it is only reasons for incompatible actions that have relative weight. This is a strange outcome, since intuitively reasons for a compatible alternative action, and even reasons for the same action, can differ in strength. For instance, the fact that I promised to push the red button is a reason to do so, and the fact that five people will die unless I push the red button is also a reason to so, and the latter is the stronger reason. Scanlon's reply to this is that in order to determine the relative weights of these two reasons, we would need to look at a counterfactual where c' is as similar to c as possible but the two facts are reasons for incompatible actions. A problem with this approach, as (Sampson, 2015) has pointed out, is that we can imagine pairs of reasons for which it fails: Either it is impossible for both to obtain in the same set of circumstances or it turns out that in the set of possible circumstances where they do both obtain we get the wrong result. I shall return to Scanlon's view of outweighing on pages 54-55 in the next chapter

For Skorupski, on the other hand, a *pro tanto* reason's possession of a certain degree of strength is seen as an intrinsic aspect of the normative property of being a *pro tanto* reason. Given this, it seems that Skorupski will have access to the resources necessary to understand strict non-weighted reason relations in terms of weighted reason relations. On Skorupski's view however, this is not the case (2010a, pp. 39 - 41).[30]

The explanation of why he holds this view is as follows. When we are talking about reasons for beliefs it is sometimes not true that we have sufficient reason to hold the belief that p even though we have stronger reasons to believe that p than we do to believe that not-p. This is the case, according to Skorupski, when, for example, the reasons for believing p are very weak but still stronger than the

---

[30] It should be noted that Skorupski does not make the distinction between sufficiency and conclusiveness; he makes do with sufficient reasons. Still, I believe it is safe to assume that Skorupski *would* deny both my definition of sufficient reason and my definition of conclusive reason, since they both are cast in terms of weighted notions of reasons. For Skorupski, having a sufficient reason does not entail that there is more reason to do that action than any other; rather it is a reason that is "'good enough' for action" and "We say that there is *sufficient* reason to spend the afternoon at the cinema, without taking that action to be optimal" (Skorupski, 2010a, p. 39).

reasons not to believe *p*. In such a scenario we should withhold judgment as regards *p*.

When it comes to reasons for actions things might be different for Skorupski. This is because in the case of actions everything but *φ*ing is equivalent to not-*φ*ing – there is no third option, the option corresponding to withholding belief. The notions of a conclusive reason and a sufficient reason are supposed to be orthogonal to what type of reasons we are dealing with, be it theoretical, practical, or evaluative. Since we cannot understand the strict notion of sufficient reasons in terms of the non-strict notion of *pro tanto* reasons in the epistemic case, Skorupski argues that it is not possible to fully understand these strict reason relations in terms of non-strict weighted reason relations.

Although I cannot give a complete defense here of why I think Skorupski is wrong in this instance, I can at least sketch the argument. Skorupski denies that the strict notion of reason can be understood in terms of weighted ones because, in the epistemic case, we should sometimes suspend judgement, and believe neither *p* nor not-*p*, even if the reasons in favor of believing *p* outweigh those in favor of believing not-*p*. However, it is quite possible that it is never the case that we should withhold belief. Rather, we should hold beliefs to a certain degree. So, in the case where we have weak reasons to believe *p* but even weaker reasons to believe not-*p*, instead of withholding judgement we should believe *p* to degree >0.5 and not-*p* to degree <0.5. The degrees of belief should, in the ideal scenario, add up to 1. A proper defense of this claim would take us to far away from the central issues in this dissertation, so I will leave it as a rough outline of an alternative to Skorupski's picture.[31]

Overall Reasons

I want now to turn to a reason relation that is often contrasted with *pro tanto* reasons, namely overall reasons. I shall characterize the reason relation and argue in favor of the usefulness and reasonableness of including overall reasons in our reasons-discourse.

---

[31] For a general overview of credence and degrees of belief, see (Buchak, 2014).

While we often talk about *pro tanto* reasons to do things, sometimes we also make judgements about what we have overall reason to do.[32] A useful metaphor is that of making a list of pros and cons with regard to whether to φ or not-φ. We have a pro-column and a con-column. We can ask ourselves what each of the individual pro or con items amounts to, but we can also ask what all of the items in the pro-column *taken together* amount to, and what all of the items in the con-column amount to *taken together*. The items in each column, taken together, correspond to the overall reason for φing and the overall reason against φing.[33]

To illustrate, take a member of parliament who has two options, voting for universal basic income or voting against it.[34] The MP has been presented with numerous reasons for each alternative by her staff, and in the end she asks:

---

[32] Other labels used in the literature to denote the concept of an overall reason include "complete reason" (Raz, 2006), "Ultimate reason" (Bedke, 2011), and "all-things-considered reason" (Chang, 2004). Overall reasons are sometimes the subject when people talk about what we have "most reason" to do. The term "overall reason" comes from (Dancy, 2004) and (Skorupski, 2010a). It is a question requiring subtle exegetical work that need not detain us here whether there might be slight differences of intended meaning to be identified in the ways the terms are employed.

[33] Then there is the question of *all-things-considered* reasons. Now, the following is not standard nomenclature (or maybe it is, it seems to be used in a variety of ways in the literature) so I will not push for my preferred understanding of *all-things-considered* reason in the main text. On my view what both columns amount to; what the entire list taken together counts in favor of is the all-things-considered reason to (not)-φ. Of course, what the all-things-considered reason counts in favor will be co-extensional with what there is conclusive reason to do. So it might be a distinction without use.

There is also another possible understanding of *all-things-considered* reason. This understanding can be exemplified by the following example: (1) "The fact that going for a walk is enjoyable is a strong *pro tanto* reason in favor of going for a walk as far as my wellbeing is concerned. It is, however, not a strong *pro tanto* reason *all-things-considered* since if I take a walk I will miss my own thesis defense." (2) "The fact that I promised is a *pro tanto* reason to do so, but it is not an *all-things-considered pro-tanto* reason to do so because I was relieved of my promise". In both (1) and (2) the *all-things-considered* refers to the circumstances of the reason, given very narrow circumstances a reason might be strong or a fact might be a reason, but given 'all the things' the reason is not strong, or even a reason at all. For more on all-things-consideredness see (Andersson, 2017, pp. 161-162).

[34] In the example, I assume for simplicity's sake that it is not possible for the MP to abstain from voting. Adding a third option does not change anything structurally. It merely adds to the length of the example (even more so than this footnote, including this parenthesis).

"(Overall), how strong are the reasons to vote in favor of the motion, and (overall), how strong are the reasons to vote against the motion?"

What she is asking for is not an additional *pro tanto* reason, but rather what all of the *pro tanto* reasons in favor of voting for universal basic income amount to when taken together, and likewise what all of the *pro tanto* reasons against voting for universal basic income amount to.

One of her staffers might respond to the MP's question by saying: "The overall reason for voting in favor of the motion is very strong, stronger than the overall reason to vote against it!" The staffer might be correct even if it is the case that each specific *pro tanto* reason to vote for the motion on its own is quite weak and each specific *pro tanto* reason to vote against the motion is quite strong. More generally, *pro tanto* reasons can be individually weak but when taken together as an overall reason have a totally different strength that is not identical to any specific *pro tanto* reason.[35] Instead of 'taken together', we could talk of 'the combined *pro tanto* reasons' or 'the *pro tanto* reasons collectively'. I am not attached to any specific way of phrasing it. What I am attached to is the view that the notion of 'overall reasons' is not only intelligible but also plays an important role in the structure of normativity. This is not an uncontroversial position (on the other hand, it is not an uncommon position either) and we will analyze the objections that have been raised against it as soon as the concept has been properly introduced.

The overall reason-relation has the following structure:

Overall Reason: OR(x, S, w, φ)

A fact (or collection of facts), x, is an overall reason for an agent, S, with a certain degree of strength, w, to φ. The fact, x, is constituted by *pro tanto* reasons to φ (and only those to φ, not by the reasons not to φ). To illustrate, take the example of the MP and her *pro tanto* reasons for voting for the motion. Assume that there are two reasons in favor of voting for universal basic income, x_1 and x_2. Let [x_1, x_2] denote a complex reasons-fact, i.e. a fact taking two (or more) reasons together. In our imagined situation the overall reason to vote in favor of universal basic

---

[35] As we will see in the chapter on the weighing metaphor, it is not only that the strength of the overall reason need not be identical to that of any specific pro tanto reason, but also that the strength might not even be reducible to the sum of the strengths of all pro tanto reasons.

income is the complex fact $[x_1, x_2]$, which is composed of the two *pro tanto* reasons in favor of voting for universal basic income. To simplify the notation, I will refer to the complex fact $[x_1, x_2]$, i.e. the overall reason, as o_1. The strength of the overall reason can differ from that of any individual pro tanto reason it is constituted by. However, the overall reason for S to φ can only be constituted by *pro tanto* reasons that are *pro tanto* reasons for S to φ. For instance, there is no overall reason for the MP to φ constituted by the MP's *pro tanto* reason, x, to φ and David's *pro tanto* reason, y, to ψ.

In general, the overall reason for S to φ is to be understood as all of the *pro tanto* reasons for S to φ taken together. Overall reasons, as has been pointed out by Brown (2014), involve collective plural predication of the form "The Fs are G". Borrowing Brown's example to illustrate collective plural predication:

(1) The pencils are five inches long.

(2) The pencils are scattered around the room (Brown, 2014, p. 2).

In (1) the property of being five inches long is attributed to each pencil individually. Every single pencil is five inches long. In (2), however, none of the individual pencils is, on its own, scattered across the room. It is only the pencils, taken together, that are scattered around the room. The property of being scattered is ascribed to the fusion of the pencils. In (1) we are dealing with distributive plural prediction, and in (2) we are dealing with collective, or non-distributive, plural predication.

Analogously, we can make the same maneuver with reasons. A slight annoyance, however, is that, unlike "being scattered", the property of having a certain strength, such as "being very strong", has both a distributive reading and a collective reading. When we want to signal that the collective reading is intended we add the modifier '*overall*'. Compare the following two sentences.

(3) The reasons to φ are very strong (each *pro tanto* reason is very strong).

(4) Overall, the reasons to φ are very strong (the *pro tanto* reasons taken together are very strong, i.e. the overall reason to φ is very strong).

In the example of whether the MP should vote in favor of universal basic income or against it, there are two overall reasons. There is one overall reason to vote in

favor of the proposal, constituted by the *pro tan*to reasons to vote in favor, and one overall reason to vote against the proposal, comprising the *pro tanto* reasons to vote against the motion. A phrase often used in discussions of this kind is "the balance of reasons". On my proposal the balance of reasons favors is to be identified with the strongest overall reason. "The strongest" overall reason is, of course, just another way of saying the conclusive overall reason. For instance, if the overall reason to vote in favor of the proposal is stronger than the overall reason to vote against it, the balance of reasons favors voting in favor of the proposal.

Some have argued that the existence of overall reasons would lead to a vicious regress – continuously changing the balance of reasons. Dancy writes:

> Much of our talk of reasons is about contributory reasons in this sense, reasons on one side or on the other, reasons that stack up with others to make a better or worse case for an action. But as well as talking about reasons in this way, we also speak of what there is overall reason to do. There is nothing wrong with that of course, but it should not delude us into thinking that there are such things as overall reasons in addition to contributory ones. […] We can say we have more reason to do this than that, but most reason to do some third thing. These verdicts do not themselves specify further reasons (of an overall sort), on pain of changing the very situation on which they pass verdict. So there are no overall reasons (Dancy, 2004, p. 16).

These remarks of Dancy's can be read in two different ways. First, he says that there are no overall reasons "*in addition* to contributory ones". He is in one sense correct – an overall reason is not a reason over and above the pro tanto reasons taken together. It is not an additional *pro tanto* reason that should be taken together with the other *pro tanto* reasons. However, Dancy goes on to say that "So there are no overall reasons". If all Dancy intends to show is that overall reasons are not *additional* reasons,

I agree with him. For the argument's sake, however, suppose that he believes, as he says later in the passage, that there are no overall reasons. Dancy's argument for the rejection of overall reasons can be rephrased as follows: To be a reason is to count in favor of something, and therefore to be something that should be given weight in our deliberation. Assume that we have two *pro tanto* reasons, x_1 and x_2, and that both are reasons to φ (rather than not-φ). We weigh these up and reach a verdict. If by taking these two reasons together we can get an overall reason o_1, then, all of a sudden, there is a new reason. Since o_1 was not included in our

initial verdict, our verdict is undermined, or in Dancy's words "the very situation on which we passed verdict has changed" and we would need to start all over and weigh up x_1, x_2 and o_1 to reach a new verdict. If there are such things as overall reasons weighing up the three reasons would result in yet another overall reason, o_2, and so it would go on *ad infinitum*, never reaching a final verdict, with the normativity of the situation ever-changing with the addition of new reasons. It seems hard to deny that if there is such a thing as overall reasons, then they have the property of counting in favor of something. It also seems reasonable to insist that they should be given weight in our deliberation.

There is a fundamental problem with this reconstructed Dancyean argument against the existence of overall reasons. Brown (2014), as far as I am aware, was the first to point out this problem. My approach is a bit different from Brown's, but it hinges on the same insight. Brown interprets reasons as properties, and hence he has the tools of mereology to work with. Given that within mereology whole-formation is idempotent (i.e. the fusion of a thing with the thing itself just is the thing itself), he can show that no vicious regress is generated by overall reasons. The conception of reasons as facts that has been assumed throughout this thesis makes things a bit different.

Roughly speaking, there are two families of views about facts: A worldly view of facts and a conceptual view of facts (Correia, 2010, p. 258). In order for the argument that I am about to present to work we need to assume a worldly view of facts, or we need to take [P] and [P & P] to express the same fact. To get an idea of how the worldly and conceptual view of facts differ, consider the following examples from Correia (2010, p. 259):

- [a is a water molecule]; [a is an H_2O molecule];

- [France is east of Argentina]; [Argentina is west of France];

- [France is east of Argentina]; [it is not the case that it is not the case that France is east of Argentina].

According to Correia, on a worldly conception of facts "each list is constituted by one fact appearing twice", but on a conceptual conception of facts "each list will naturally be taken to compromise two distinct facts" (Correia, 2010, p. 259). Whether each line on the list is to be considered as two facts depends on whether

your conception of facts tracks what is out in the world or distinct propositions. Thus [France is east of Argentina] and [Argentina is west of France] pick out the same thing in the world – the propositions have the same truth-makers – but they are, on most views, two distinct propositions. For my argument to work, we need to take [P]; [P, P] to express the same fact rather than two distinct facts.

I earlier defined an overall reason to φ as a complex fact consisting of the *pro tanto* reasons-facts to φ [x_1, x_2]. With this in mind, let us rephrase Dancy's argument and show that the regress he is anxious about never gets off the ground, or that it is at least not vicious.

1. $O_1 = [x_1, x_2]$. (Assumption, the overall reason is the complex fact consisting of pro tanto reasons)

2. $O_2 = [[x_1, x_2], O_1]]$. (Assumption, Dancy is correct about the regress)

3. $O_2 = [[x_1, x_2], [x_1, x_2]]$ (From 1, 2).

4. $[[x_1, x_2], [x_1, x_2]] = [x_1, x_2]$ (Assumption, that [P] and [P, P] express the same fact)

5. $O_2 = [x_1, x_2]$ (From 3, 4).

6. $O_1 = O_2$ (From 1, 5)

As we can see, the regress is no more vicious than the regress of identity, i.e. no more vicious than saying that A=A … A=A. The overall reason does not change the balance of reasons. Of course, on this view overall reasons have no normative strength independently of the *pro tanto* reasons they are constituted by. Dancy's problem stems from the assumption that when we weigh our *pro tanto* reasons together and arrive at an overall reason, the overall reason has the combined force of the *pro tanto* reasons, and that this combined force should be added together with the *pro tanto* reasons. This is akin to believing the following: If we have two pens each of which individually weighs 10 grams, when we weigh these together we get 20 grams. Then, when we weigh the individual pens and the collection of pens at the same time, the resulting weight is 40 grams. But weighing all of the individual pens, on one hand, and weighing the collection of pens, on the other,

is the same thing: In both cases the scale will say 20 grams.[36] It might be argued at this juncture that an overall reason fails a requirement for being a reason because it fails to count in favor of something, but this is mistaken. It does count in favor of something, just as the collection of two pens does have a weight. It is merely that the favoring in question is nothing over and above the favoring of the *pro tanto* reasons. Likewise, the pen collection has no weight over and above that of the individual pens.[37]

Furthermore, there is something peculiar in Dancy rejecting overall reasons but also saying that *pro tanto* reasons make a better or worse *case* for acting (as he does in the passage I quoted above). His talk of something making a case stronger or weaker indicates that it is at least possible, and maybe even plausible, to understand Dancy's phrase "an overall case for action" as denoting the very same concept that I denote with "overall reason for acting". One reason to believe that this is the case is noted by Skorupski, who writes:

> However, … [Dancy's] point, I believe, is that overall reasons do not constitute *additional* reasons in the sense that they cannot change, so to speak, the reason-giving situation determined by the totality of [*pro tanto*] reasons (Skorupski, 2010a, p. 40).

And indeed, as I have shown, just as the pens taken together do not give any weight *in addition* that of the individual pens, neither is the overall reason a reason that has normative strength *in addition* to the *pro tanto* reasons. The analogy that I have made with pens and reasons, however, only stretches so far. In a contrast with weight, it is far from clear – it is perhaps even implausible to suppose – that the normative strength of *pro tanto* reasons taken together is a strictly additive function.

I end by making a few concluding remarks about overall reasons. The overall reason for S to φ can only be constituted by S's *pro tanto* reasons to φ. It cannot be partly constituted by *pro tanto* reasons for another agent to φ, or by *pro tanto* reasons to do an entirely different action, ψ. All of this seems quite straightforward

---

[36] A similar question, eagerly discussed in the debate over collective agency, is whether, if two agents go for a walk together, there are two or three agents walking (the first person, the second person and the collective agent). For a discussion of this, see (Gilbert, 1990).

[37] For more on 'over and above-ness' see (Smid, 2017).

and intuitive. Some might, however, wish to hold that the overall reason to φ is constituted not only by the *pro tanto* reasons to φ, but also by the *pro tanto* reasons to not-φ – that is, by both the pro-column and the con-column. I think this is wrong for one simple reason: The overall reason to φ is supposed to be everything that counts in favor of φing and the *pro tanto* reasons against φing cannot be a part of what overall counts in favor of φing. The fact that I will break my promise if I save a child that is about to drown does in no way in count in favor of saving the child. It is not *because* the reasons against doing something I have a reason to do it, it is despite of it.[38]

As a preliminary and a way into the discussion, presented in the chapter "On Buck-Passing", on the connection between what we have reasons to do and what we ought to do, let me now state my view: What we ought to do is determined by, or grounded in, the balance of reasons we have. *S* ought to φ if and only if, and because, *S* has a conclusive reason to φ.

---

[38] However, if we have a *conclusive* overall or *pro tanto* reason to φ the fact has the property of being *a conclusive reason* partly in virtue of the reasons against φing. This is because it is a conclusive reason partly in virtue of being stronger than the reasons against φing and in order for this relation to obtain the reasons against needs to play a part. Compare this with footnote 32, p. 18, it might make more sense to say that I have an *all-things-considered* reason to φ partly in virtue of the reason against φing.

The Weighing Metaphor

In the previous chapter we touched upon the idea that *pro tanto* reasons and overall reasons have a degree of strength or weight.[39] In other words, reasons count in favor of an option to a certain degree. In this chapter we focus on the comparison of reasons with each other, and on how to rank reasons in accordance with their strength and how to weigh them. I will begin by comparing analogous normative weighing and ordinary mechanical weighing. I will then go on to discuss some of the properties different possible weighing functions might have. Ultimately, I will argue that the idea of weighing, as in making pair-wise comparisons of the strength of different reasons, is relatively unproblematic. What is problematic is how to determine the weight of individual *pro tanto* reasons and overall reasons.

I dare say there are few fields in philosophy so firmly in the clutches of a metaphor as the field of reasons and normativity. The metaphor in question is the *weighing metaphor*. This metaphor is meant to illustrate the idea that reasons are like weights placed on a scale, with reasons for doing something placed on the one side and reasons against doing that thing placed on the other side. The way the scale tips determines what you have the weightiest, or strongest, reason to do. Talking about reasons and their associated weights in terms of mechanical weighing is meant to be helpful and illuminating.

The weighing metaphor is without a doubt helpful in that it provides us with a vernacular for talking about reasons, and the fact that some count for more with us than others.[40] There is, however, a worry that we could interpret the metaphor

---

[39] Just to remind the reader, I use "the strength of a reason" and "the weight of a reason" and associated phrases interchangeably.

[40] The weighing metaphor is ancient. For example, representations of the scales of justice, the balancing of interests and evidence, can be traced back at least to ideas about the ancient Egyptian goddess Maat and Greek Lady Justitia.

too literally and believe that weighing reasons is identical to normal mechanical weighing.

Take the following example of the weighing metaphor being put to use:

> Alexander Hamilton promised Aaron Burr to meet Burr in the room where it happens. On Hamilton's way to the room where it happens Hamilton encounters a child that is about to drown. Hamilton can save the child at the expense of breaking his promise to Burr. In this scenario, intuitively, the fact that Hamilton promised Burr is a *pro tanto* reason that carries a certain weight for keeping the promise and thereby letting the child drown. Likewise the fact that the child is about to drown is a reason for Hamilton with a certain weight to save the child and thereby break his promise to Burr. Hamilton weighs his reasons and decides that the reason to save the child to outweigh his reason to keep his promise to Burr.

On the weighing metaphor, these two reasons are to be put on opposite sides of a scale. Assuming that the scale tips in favor of saving the child, the reason to save the child is weightier or stronger than the reason to keep the promise.

All of which seems simple enough. There are, however, certain pitfalls that we must be aware of when employing the weighing metaphor.

Mechanical Weighing and Normative Weighing

The picture that the weighing metaphor supplies us with is one suggesting that, in principle, we should consider the weighing of reasons as a process that is somewhat analogous to mechanical weighing. Mechanical weighing, however, seem to differ from the weighing of reasons in some key aspects. An important difference is that mechanical weighing not only conforms to an ordinal scale but also to a ratio scale. A 100kg gold bar is not only heavier than a 50kg gold bar – it is twice as heavy. Additionally, we can assign a specific number and unit of measurement to the gold bars: The first is 100kg. In other words, in the context of mechanical weighing we can apply both a ratio scale and an interval scale, and we can assign more or less precise numbers to the objects being weighed.

When it comes to weighing reasons (and other gradable normative phenomena such as values) things might not be so easy. In particular, it is doubtful that we can sensibly put a numerical number on the weight of a reason. Broome writes

the following about numerical values and the weight of reasons: "[T]he weight of a reason will rarely be such a precise thing as a number. It is likely to be an entity of some less determinate sort. Weights will not therefore combine by simple addition" (2013, p. 52).

I believe that we can partly side-step the problem of being unable to ascribe a number to a reason's weight. Even if we cannot assign a specific numerical value to the weight of a reason, I take it as given that we have *some* understanding of how weighty specific reasons are. For example, I think few would deny that a reason to save a child from drowning outweighs a reason to keep a promise to water some plants. Furthermore, in some cases it even seems to be accurate to say that we can represent the weights of reasons with numerical values. We can imagine a situation where the only difference between two choice options is how much welfare we will gain. In most of theories of the value of welfare, if one option will supply us with twice as much welfare as the other, we can reasonably say that the reason for taking the first option is twice as strong as the reason for taking the second.[41] In other words, we seem to be able to ascribe a ratio between the weights of the two reasons.

Not all cases are like the one just mentioned. Thus, while it seems correct to say that the child-saving-reason is stronger, or even a lot stronger, than the promise-reason, it seems a lot harder to assign a ratio between the two. Is the child-saving reason twice as weighty, or four times as weighty, as the promise-keeping reason?

Another breakdown in the metaphor is that it is not obvious that all reasons are comparable in terms of their weights. In the case of mechanical weighing, regardless of what we put on each side of the scales we will reach a definite conclusion about whether what is on the right side of the scale outweighs, is the same weight as, or is outweighed by, whatever is on the left side of the scale. Mechanical weighing is in this sense fully commensurable. There are no comparisons of two objects (with mass) in terms of weight that are not, in principle, determinable.

Furthermore, in the case of mechanical weighing it is, at least in principle, always possible to add enough weight on one side of the scale to tip the scale. Where the weighing of reasons is concerned, this can be questioned. A Millian or a Kantian

---

[41] Whether we can assign ratios to welfare can, however, be questioned. For more on this, see (Harsanyi, 1986; Broome, 1991, 2004b).

might want to say that there are some reasons that could not, even in principle, be outweighed by another reason (or collection of reasons). Some reasons have superior weight.[42] While these first-order theories may be incorrect, we should not rule them out on conceptual grounds. Therefore, our understanding of the weighing metaphor in the normative domain should allow for the conceptual possibility of a reason with superior strength: The weight of reasons might not have the Archimedean property.[43]

In mechanical weighing, we never reach a situation where it is indeterminate which side weighs more. In addition, if the left pan of the scale outweighs the right, any added weight added to the right pan will make the weight difference smaller. The same is not necessarily true when reasons are weighed.

Take the following classic example of a hard choice.

> *Hard Choice*: M can choose to pursue either a career in law or a career in philosophy.

The topic of whether all objects stand in one of the three positive value relations – i.e. better than, equally as good as, or worse than – has received a lot of attention in recent years (Parfit, 1984; Raz, 1986; Chang, 2002a; Broome, 2004b; Chang, 2004; Rabinowicz, 2008, 2012; Chang, 2013; Elson, 2014; Parfit, 2016; Andersson, 2017; Elson, 2017). What we are interested in here is the similar question: For any reasons *A* and *B*, is it the case either that *A* is a stronger than *B*, or that *A* and *B* are equally strong, or that *A* is weaker than *B*? Another way to put it is: In the context of *Hard Choice*, are the reasons for M to choose to pursue a career in law stronger than, of the same strength as, or weaker than, her reasons to pursue a career in philosophy?

In *Hard Choice* it seems that the reasons for a career in law are neither stronger nor weaker than the reasons for a career in philosophy. One might therefore be

---

[42] For more on superiority, and especially superiority in value, see (Arrhenius & Rabinowicz, 2015).

[43] The Archimedean property is commonly used in mathematics – e.g. the Archimedean property of real numbers can be described as: "For any real numbers x and y there is an n such that $nx \geq y$" (Arrhenius & Rabinowicz, 2015).

tempted to conclude that the reasons are equally strong. However, consider *Hard Choice**

> *Hard Choice**: M can choose to pursue either a career in law or a career in philosophy. If she chooses to pursue a career in law, she gets a cookie.

The idea with *Hard Choice** is that it is identical with *Hard Choice* except insofar as there is a small improvement to one of the options.[44] If the reasons for both careers are equally strong, this improvement ought to ensure that the career in law is favored by the balance of reasons. Another possibility is, of course, that the cookie does not improve the option of having a career in law. If it is insisted that whether or not we get a cookie is not a reason for picking a career, we can easily replace the cookie with some other small improvement.

We will not dwell too long on this issue. I will not try to settle whether there is such a thing as incomparability, and I will not attempt to establish how best to account for incomparability, be that with vagueness (Andersson, 2017), a fourth or fifth value relation (Chang, 2002a; Rabinowicz, 2008, 2012) or some other tool.[45] What I want to stress is that our understanding of the weighing metaphor should not be taken too literally so that it rules out the possibility of incomparability of reasons.

Atomism or Holism

I have so far pointed out some breakdowns in the weighing metaphor. Next I wish to focus on the process of weighing reasons. In particular, I want to ask whether that process is additive. Mechanical weighing is a monotonic and additive process. If we place two 100kg gold bars on the same side of the scale there will be 200kg of gold on that side of the scale. The weight of any specific gold bar does not depend on any other gold bar. In the passage excerpted above on page 43, Broome claims that the weighing of reasons does not work in a similar fashion. The weight

---

[44] This is a version of Ruth Chang's small-improvement-argument. For more on this, see (Chang, 2002b; Gustafsson, 2013; Andersson, 2017).

[45] For an overview and more on value relations, incomparability, and hard choices, see (Chang, 2002a; Andersson, 2017).

of reasons is not representable by something as fine grained as a number and "weights will not therefore combine by simple addition" (2013, p. 52).

I think that Broome is correct that the unavailability of precise numbers entails that the process of weighing reasons does not allow *simple* addition. However, it is still of interest to discuss whether weighing reasons is still additive in a *less* simple sense. Regardless of whether we can ascribe a specific number to the weight of a reason, it may be that the process is additive in the following sense: The weight of two reasons taken together is a monotonic and strictly increasing process and is the 'additive' sum of the weight of the two reasons taken in isolation. We can see the appeal of the idea that reason weights are additive. It makes the process of weighing simpler and more straightforward.

In order to have additivity in the sense that the process of weighing reasons increases monotonically, the weight of reasons must be such that that for any *pro tanto* reasons x_1 to φ with strength w_1 and x_2 to φ with strength w_2 the total strength, $w_{1,2}$, of [x_1 and x_2] must be more than the strength of either x_1 and x_2 alone.[46] To meet this condition, we must assume atomism about reasons.

By atomism about reasons I mean that reasons are context-insensitive.

> *Atomism*: If x is a reason to φ in circumstance c then x is a reason to φ in any circumstances $c*$.[47]

Strictly speaking, we need a stronger version of atomism in order to achieve additivity. Atomism as stated here says merely that a reason's *valence* is invariable, or context-insensitive. By "valence" I mean the normative direction of a reason – i.e. its being *for φ* or *against φ*. What we need is *strong atomism*.

Strong atomism says that reasons retain not just their valence regardless of circumstances, but also their specific weights: The weight of a reason is context-insensitive.

---

[46] For present purposes we do not require the sum to be representable numerically.

[47] For more on different formulations of atomism about reasons see (Dancy, 2004, p. 74; McKeever & Ridge, 2005)).

> *Strong Atomism*: If *x* is a reason to *φ* with a certain weight, *w*, in circumstance *c*, then *x* is a reason to *φ* with at least weight, *w*, in any circumstances *c**.

The negation of atomism is *Holism*.

By "holism about reasons" I mean the position that reasons are context-sensitive.

> *Holism*: Even if *x* is a reason to *φ* in circumstance *c* this does not entail that *x* is a reason to *φ* in any circumstances, *c**. In *c**, *x* might even be a reason to not-*φ* (Dancy, 2004, p. 74).

Just as with atomism, we can distinguish between a weak and a strong version of holism. Weak holism is enough to preclude additivity. It simply is the negation of strong atomism.

> *Weak Holism*: If *x* is a reason to *φ* with a certain weight, *w*, in circumstance *c*, then *x* is a reason to *φ* in any circumstances *c**, but *x* but in *c** it is possible that *x* will have less weight than *w*.

Our weighing function will not be additive if weak holism is true. Since holism entails weak holism, we also lose additivity if holism is true. We lose it because it will no longer be necessarily true that the weight of two reasons taken together is the sum of the weight of each reason taken on its own, and because the addition of a third reason will not necessarily make the resulting combined reason have a weight equivalent to that of the previous two reasons plus the weight of the new one. A change in the circumstances, such as the addition of a third reason, might change the weight of the other reasons, and the sum of the weight of the reasons in this circumstance would differ from the sum of the weight that each reason has on its own.

In addition to the two versions of holism and the two versions of atomism just mentioned, there is a wide array of different versions of holism and atomism, including atomism and holism in value theory. For present purposes we can bracket these complications.[48]

---

[48] For instance, we bracket whether *very strong holism*, which claims that no reasons are context-independent, is valid or not. We also bracket discussion of whether there is a connection between holism about reasons and the debate on *generalism* and *particularism* about moral principles. Roughly, generalism is the view that there is a general moral principle which is

One of the more popular arguments in favor of holism invokes the phenomena of background conditions. Exactly what background conditions are, and how they should be understood, is a topic much discussed. In Paper I of this dissertation, my co-author Andrés Garcia and I go into more detail on how we believe background conditions should be understood. I will therefore not go into great detail in this introduction. Instead I will merely give an intuitive sense of what background conditions are, and what explanatory role they supposedly have. Roughly speaking, there are four kinds of background condition. Dancy calls them *enablers, disablers, intensifiers, and attenuators* (Dancy, 2004). An enabler is a feature that needs to be present in order for the reason, value, ought or other normative phenomenon to obtain. A disabler is a feature the presence of which undercuts normative phenomena – e.g. it may be that a fact *x* with the property of being a reason loses that property when a disabler is introduced and fails to have that property if the disabler is already present. An intensifier is a feature the presence of which makes a reason stronger, a value more valuable, and so forth. An attenuator is the opposite of an intensifier, i.e. a feature the presence of which makes a reason weaker.

It is important to note that the background conditions are not a part of the reason, or the value. To see this, think about the following example: In general, the fact that I promised my friend to do something is a reason to do it. The reason-fact in this scenario is the fact that I promised. However, there seem to be quite a few features that need to be in place in order for the reason-fact to have the property of being a reason. One such feature is the fact that my friend has not released me from my promised. If my friend were to do so, the fact that I have promised would no longer be a reason for me to do it. In this scenario, the fact that I have *not* been released of my promise functions as an enabler. But that fact is not a part of my reason to fulfill the promise. It is not part of the reason-fact.

If I am not released from my promise, on the other hand, my promise is a reason. We can suppose that I have a history of not keeping my promises. This fact might make the fulfilment of the particular promise especially important. In other words, my history of breaking promises might serve as an intensifier – making my

---

applicable to all and every situation, and particularism is the denial of generalism. For more on these debates, see (Crisp, 2000a; Dancy, 2000; Jackson, Pettit, & Smith, 2000; McNaughton & Rawling, 2000; Raz, 2000; Dancy, 2003; McKeever & Ridge, 2005, 2006; Raz, 2006; McKeever & Ridge, 2013; Garcia, 2018a).

reason to keep the promise stronger than it otherwise would have been. Neither the fact that my friend has not released me from my promise nor the fact that I have a history of breaking promises is a part of my reason to keep the promise. They are auxiliary features and not parts of the reason-fact. The fact that has the property of being a reason is the fact that I promised. The auxiliary features function as reason-makers without themselves being parts of the reason-fact.

According to weak atomism, if a fact is a reason in one set of circumstances, then it is a reason in any set of circumstances. It follows that if there are such as things as enablers and disablers, weak atomism is false – false, because the enablers and disablers would ensure that it is not necessarily true that a fact that has the property of being a reason retains that property across all circumstances.

According to strong atomism, if a fact is a reason with a certain degree of strength in one set of circumstances, then it is a reason with the same degree of strength in any set of circumstances. Therefore, even if there are no enablers or disablers, but there are such things as intensifiers and attenuators, then strong atomism is false.

Proponents of atomism therefore need to deny the existence of background conditions. However, this is not easy to do. The denial goes against our common-sense intuitions about cases such as the one involving promising. The most common strategy adopted by those who reject background conditions is to expand the reason-fact.[49] According to such a deflationary view of background conditions, the fact that I have a history of breaking promises and that my friend has not released me from my promise is a part of the reason-fact.

In order to account for our intuitions about these cases, the expansion of the reason-fact, i.e. what actually constitutes the reason, is often accompanied by a pragmatic story. The story is needed to explain away the intuitions we have – e.g. the intuition that the fact that my friend has not released me from my promise is no part of my reason to keep it.

The pragmatic story comes in various guises, but all versions deploy a distinction between everyday use of the term 'reason' and a technical, complex use of the term 'reason'. In general, when we want to explain something to someone, such as what reasons they have, what we treat as *the* explanation is dependent on communicative norms. For instance, if someone were to ask me why a plane

---

[49] See (Crisp, 2006; Raz, 2011; Fogal, 2016).

crashed, I might present a story about pilot error or some technical fault. I would not mention gravity or laws of aerodynamics. Of course, gravity and the laws of aerodynamics do play a part in a complete explanation of why the plane crashed, but I presume that the listener already knows about gravity and has a rough enough understanding of aerodynamics. The same happens when we identify as reasons. In explaining what the reason is, and why we have the reason, we tend – to borrow Daniel Fogal's phrase – to pick out a good representative of the complex reason (Fogal, 2016, p. 92). We only mention the fact which, in the communicative context we find ourselves in, is the most effective way of communicating the presence of the complex reason. To illustrate: In most contexts, it would not be communicatively virtuous to mention that my friend has not released me from my promise. If I say that I have promised to do it, it is implicit that my friend has not released me from the promise. As has been pointed out to me by Jonas Olson, understanding reasons as these expansive 'totality facts' makes it hard to see how one would understand the process of *weighing* in the first place, because on this picture the weighing of reasons seems to become a matter of mere pragmatics.

An issue discussed in more detail in the paper on background conditions is the following: Even if we agree that our identification of reasons in everyday communication is to some extent governed by communicative norms, this, in itself, is compatible with the idea that there is a distinction to be made between reasons and background conditions in the technical sense of the term "reasons".

Another common argument against atomism – and *a fortiori* against strong atomism – is the argument from *malicious pleasures*. So, it might be prudent to run through the familiar example here. McKeever and Ridge describe it in the following way:

> *Malicious Pleasure*: Suppose my seeing a movie would give me pleasure. Plausibly, the fact that seeing a movie would give me pleasure is a reason for me to see it. However, in another case that very same consideration might be no reason at all to perform the action. If, for example, I would take pleasure in seeing someone tortured in the movie (suppose it is a documentary on human rights abuses) then the fact that I would take pleasure in the film plausibly is no reason whatsoever to see it (McKeever & Ridge, 2006, p. 27).

As you can see, this example hinges on our intuitions about background conditions. Pleasure is normally taken to be a paradigmatic contender for something that there is always a reason to seek. In the first example in the passage, this may seem quite plausible. In the second part of the example, however, involving taking pleasure in seeing someone tortured, it does not seem plausible anymore. The fact that going to the movies would cause me pleasure no longer seems to be a reason for me to go and see the movie. It even seems to be a reason against going to see the movie.

Atomists can once again reply by expanding the reason-fact and providing a pragmatic story. The proper description of any pleasure-based reason is perhaps that "it would bring about non-malicious pleasure". The fact that going to the movie would give the agent malicious pleasure would therefore not be a reason to go to the movies. There seems, however, to be a never-ending supply of examples which force the atomist to go on and expand the reason-fact *ad absurdum*. To conclude, I believe that was has been said in this discussion lends to support to the view that atomism is less plausible than some sort of holism. However, more needs to be said on the topic than I have said here.

Weight or Weightier-than

Moving on, another interesting issue concerning the weighing of reasons is whether the comparative relation of one reason being *weightier than* another is more metaphysically fundamental than the monadic property of a reason having a certain *weight*. In other words, can I have a weighty reason to φ even if I could not (even in principle) have other reasons to φ or reasons against φing?

Errol Lord and Barry Maguire characterize the two positions as follows:

> Those that take *weight* to be more fundamental think that we can analyze what it takes for particular reasons to have strength of a certain magnitude in isolation from the strengths of the other reasons. Once we determine what weight each individual reason has, we can come up with an ordering of these weights. This ordering will determine the relational *weightier than* facts. […] Those that take *weightier than* to be more fundamental think that the ordering of relative weights come first. They think that we first analyze what it is for some reason (or set of reasons) to be weightier than another reason (or set of reasons). Once we do that,

we can determine which weights individual reasons have (Lord & Maguire, 2016, p. 14).

The issue is whether a particular reason being weightier than another particular reason is metaphysically prior to the weight of the reasons in question, or vice versa.[50]

Lord and Maguire assert that it is more natural for a weight fundamentalist to hold that reasons are not fundamental, but rather understood in terms of some other notion (Lord & Maguire, 2016, p. 15), but of course this need not be the case (c.f. (Skorupski, 2010a)). Lord and Maguire give two examples: Desire-based theories of reasons and value-based theories of reasons. On a desire-based view of reasons (often referred to as "Humeanism about reasons") our reasons are understood via our desires. Since desires have strengths., a desire-based theory can quite naturally claim that the weight of a particular reason to φ is some function of the strength of the attendant desire and how well φing facilitates the desire (Lord & Maguire, 2016, p. 15).[51] On a value-based theory (where reasons are understood via values) it is likewise quite natural to claim that the weight of a particular reason to φ is some function of how much value φing would bring about.[52]

---

[50] For a debate with some overlap in value theory, please see (Gustafsson, 2014) and (Carlson, 2016) on whether the monadic property of 'being good' can be defined in terms of the comparative property of 'being better than' or vice versa.

[51] Exactly how the strength of desires should be cashed out depends on how one understands desires. Thus, if one understands desires as akin to dispositions (c.f. (Stalnaker, 1984; Ashwell, 2014)), one will interpret the strength of a desire in terms of how strong the disposition to act is. Accordingly, a desire-based approach to the strength of reasons might be able to attach a precise numerical value to the weight of reasons, thereby side-stepping some of the issues mentioned earlier.

[52] These theories are not without their own problems. Let me briefly mention a few. Desire-based theories often stall when confronted with the fact that people have desires, even strong desires, to bring about terrible states of affairs. We do not want to accept, just because of this, that we have reasons, or even strong reasons, to bring about terrible states of affairs. Value-based theories, on the other hand, often face difficulties accounting for non-axiological reasons, such as promissory reasons, where it seems that what we have reason to do would not bring about anything valuable, or even bring about something with negative value. For more on desire-based theories of reasons, see (Williams, 1981a; Smith, 1994; Schroeder, 2007; Parfit, 2011a; Manne, 2014; Scanlon, 2014).

Weightier-than fundamentalists believe that a reason *x*'s being stronger than reason *y* is metaphysically prior to reason *x*'s having a degree of strength w_1 and reason *y*'s having a degree of strength w_2. Weightier-than fundamentalists start off with the ranking of reasons, so to speak, and from that point of departure try to say something meaningful about the strength of individual reasons. There is some intuitive appeal to the idea that *weightier-than* is more metaphysically fundamental than *weight*, because reasons seem to be essentially comparative. For instance, it seems plausible to hold that when I make an assertion like "The fact that chocolate is my favorite ice-cream flavor is a strong reason for me to pick that ice-cream" I must mean that the reason is *strong* in comparison with some standard.

Normally, in a scenario where I am to choose an ice-cream flavor, the fact that a certain flavor is my favorite is a stronger reason than any other reason in play. However, if choosing that flavor has the consequence that some innocent person will die, the claim that the fact that it is my favorite flavor is a strong reason to pick it seems false. How strong a reason is – or so the intuition goes – only makes sense given a comparison class, and specifically, the position of the reason in the ranking of reasons in that comparison class. According to weightier-than fundamentalists, it does not make sense to say that a reason is, for instance, very strong (or quite weak) if no comparison is being made with the strength of another reason – assessing the strength of a reason in isolation, according to them, is meaningless.

Philosophers including Scanlon (2014) and John Horty (2012) have attempted to side-step the debate over whether a reason's weight or its being weightier-than is more metaphysically fundamental. Horty's extraordinary view on the weighing metaphor will be discussed later on pages 57-59. We can see how Scanlon rejects the dichotomy of weight and weightier-than by reminding ourselves (this was explained in the first chapter) of the way he interprets weighted reasons (e.g. *pro tanto* reasons) in terms of strict reasons (e.g. sufficient reasons).

Scanlon defines one pro tanto reason being weightier than another as follows:

> *Outweighing*: One consideration, x, outweighs another consideration, y, if the following hold: R(*x*, *S*, *c*, φ), R(*y*, *S*, *c*, ψ), where ψ is a course of action incompatible with φ, and SR(*x*, *S*, *c*, φ) but not SR(*y*, *S*, *c*, ψ), although SR(*y*, *S*, *c'*, ψ) where *c'* is a set of circumstances as normatively similar to *c* as possible except that *x* does not obtain in *c'* (2014, p. 108)

What this means is that Scanlon takes the non-weighted notion of a *sufficient reason* to be more fundamental and understands weighted reason-notions in terms of it. Both *pro tanto* reasons and overall reasons are weighted reasons, as compared with non-weighted, or strict, reasons such as sufficient reasons and conclusive reasons.[53] On Scanlon's view, in circumstances c a reason x_1 to φ outweighs another reason x_2 to ψ if x_1 is a sufficient reason in c and x_2 is not, even though x_2 would be a sufficient reason in c' where c' and c are as similar as possible except that x_1 does not obtain in c'.

A problem, raised by Maguire and Lord, is that on Scanlon's account "a reason can only outweigh its competitors when it is a sufficient reason" (2016, p. 20). The problem is that we often have reasons which, while not being sufficient, still outweigh *some* reasons. Maguire and Lord have a wonderful example of this.

> So, for example, I have some reason to take a nap right now. This reason is insufficient because I really need to get this writing done. I also have some reason to go to the airport and fly to Rome to see Bernini sculptures. This reason is also outweighed by my reason to write. It is also less weighty than my reason to take a nap – it would disrupt my life too much (Lord & Maguire, 2016, p. 20).

In this pictured scenario it seems intuitive that the reason to take a nap outweighs the reason to fly to Rome, but the reason to take a nap is not sufficient.

Scanlon's solution to this is to say that we need to find a set of circumstances c* such that, in it, only the reason to have a nap and the reason to go to Rome obtain (or at least one where one of them is a sufficient reason). The next problem is that possible worlds are tricky. It might be that in the nearest possible world where circumstances c^* obtain and so only the two (nap and Rome) reasons are in play it is not true that my reason to take a nap outweighs my reason to go to Rome. Lord & Maguire write:

---

[53] As was explained in the first chapter on reason-relations, a normative notion can be either weighted (or gradable) or non-weighted (or strict). Gradable normative notions are simply those that come in different degrees, such as *pro tanto* reasons, which come in varying degrees of strength. Strict normative notions are those that do not come in degrees, and include sufficient reasons: One cannot have a more or less strong sufficient reason – either the reason is sufficient or it is not.

> So, for example, imagine that in the nearest world where only the nap reason is in play and the Rome reason is in play, my semester is over, my writing obligations fulfilled, and my vacation impending. I'm much less tired in this world and much freer to go on vacation. In this world, it's not so clear that my reason to nap is a sufficient reason. But if it isn't, then it follows that *in the actual* world my reason to nap is not weightier than my reason to go to Rome (Lord & Maguire, 2016, p. 21).

It is surely safe to assume that the last word has not been said in this debate. I, myself, am genuinely unsure which way I lean. I do believe, however, that once we move to epistemological and first-order normative theorizing, the important question becomes whether one reason is weightier than another, regardless of whether *weigh*t is more metaphysically fundamental than *weightier-than*.

Combine, then Compare

I believe that the actual process of weighing reasons is less problematic and mysterious than is usually assumed. When "adding up reasons on one side of the scale" one does not weigh up *n*-number of separate reasons against each other. Rather, one combines the *pro tanto* reasons into overall reasons and compares the strength of the two competing overall reasons. Selim Berker (forthcoming) dubs this way of viewing the weighing process "combine, then compare". How strong the overall reasons are need not be an additive function of the strength of the *pro tanto* reasons that make up the overall reason, as an example from Horty (2012, p. 61) confirms: That it is hot outside and that it is raining are each reasons not to go for a run, but the two facts taken together is a weaker reason not to go for a run than each reason taken separately. The rain mitigates the unpleasantness of the heat and the heat mitigates the unpleasantness of the rain.

In the end, again, I take the weighing metaphor to be misleading in that it seems to allow for the possibility of weighing two or more separate reasons against one or more separate reasons. What one is doing is weighing *one* (overall or *pro tanto*) reason against another (overall or *pro tanto*) reason. It is always a case of a pair-wise comparison. Since it is a case of pair-wise comparison, the issue of additivity does not appear in the process of weighing, but only in the process of forming

overall reasons. This last claim might seem controversial, but actually I do not think it is. Consider the following case.

> *Lunch*: Ann is deciding on where to go for lunch. She can go to place A or place B. A is cheaper than B, and closer than B. But B has tastier food than A.

In *Lunch* we have three *pro tanto* reasons, *cheaper, closer*, and *tastier*.

Cheaper and *closer* are *pro tanto* reasons to lunch at A rather than B. *Tastier* is a *pro tanto* reason to lunch at B rather than A. *Cheaper* and *closer* taken together (i.e. put on the same side of the metaphorical scale) create the overall reason, [*cheaper, closer*], to go to A. Since there is only one *pro tanto* reason against going to A, *tastier*, the overall reason against going to A is [*tastier*]. Now, we do not first weigh *cheaper* against *tastier,* and then check if *closer* can make up for the remainder. Instead, we weigh [*cheaper, closer*] against [*tastier*]. We balance the weights of the two overall reasons against each other. We combine the reason-facts *cheaper* and *closer,* not the reason-properties each reason-fact has. This is not to say that we cannot weigh *cheaper* against *tastier, closer* against *tastier* or even *cheaper* against *closer*. The point is just that the ranking of *pro tanto* reasons does not necessarily tell us how the overall reasons for and against where to go for lunch compare.

When we differentiate the weighing functions from our overall reason formation principles we make the weighing metaphor look more sensible. Weighing overall reasons is easy; it is just a pair-wise comparison determining which side weighs more. Determining the weights of the overall reasons, however, is hard. I have not solved the problems of the weighing metaphor, but rather situated the problems where they should be: In the formation of overall reasons rather than in the weighing process.

Horty's Austere View

Let us now return to John Horty. He takes the weighing metaphor and the vague characterizations of weights possessed by reasons as a philosophical dead-end. "Once we move past the level of a rough outline, it will not do to say only that reasons lend some kind of weights to conclusions, and that these weights are assembled somehow" (Horty, 2012, p. 2). He then proceeds to provide a

framework that allows us to model comparisons of reasons without needing to talk explicitly about weights at all. This means we need not bother untangling the weighing metaphor.

For Horty, there is an ordering of reasons, but the ordering is not the result of a weighing procedure. Horty argues for a conception of reasons as *default rules*. A default rule is a rule which allows you to draw a conclusion, by default, when it is triggered. A default rule, or rather default reasoning, is non-monotonic. Simply put, defaults are useful heuristics, but unlike the rules of classical logic they are not always applicable. Classical logic is monotonic. What this means is that if premise A logically entails conclusion C, then you can add whatever premise, B, you want and you will still be able to derive C. In non-monotonic logics such as default reasoning there is no such security. You cannot add any premise you like and still be guaranteed that the conclusion follows. To take Horty's own example, let A stand for 'Tweety is a bird' and C stand for 'Tweety can fly' and '$\rightarrow$' stand for the default rule. Let us then define the default rule, d_1, as: 'If x is a bird then x can fly', i.e. 'A $\rightarrow$ C'. At first pass, let us define a default rule being triggered if its premise obtains (Horty, 2012, p. 25).

Given that A obtains, the default rule d_1 is triggered, so we can draw the conclusion that Tweety can fly.

Let us complicate this picture a bit. Let B stand for 'Tweety is a penguin' and d_2 stand for 'If *x* is a penguin, *x* cannot fly' or 'B $\rightarrow$ not-C'.

If both A, and B obtain, we have a problem. We have two triggered default rules, one telling us that we can draw the conclusion that Tweety can fly and another telling us that Tweety cannot fly. What we need is an ordering relation. Let '<' stand for the ordering relation of default rules. This ordering relation tells us which default rule we should follow if default rules come into conflict. Let us then assume that the ordering relation is as follows: $d_1 < d_2$.

So, while d_1 is triggered (A obtains) we know from our ordering relation that it is defeated by d_2, which is also triggered (B obtains).

This allows us to say that while Tweety is a bird is a *pro tanto* reason to believe that Tweety can fly, the fact that Tweety is a penguin is a stronger *pro tanto* reason to believe that Tweety cannot fly.

Horty's account offers an easy way to account for background conditions. We can introduce the notion of an exclusionary reason, something that triggers a default rule to the effect that one should either downplay or emphasize another default rule. This can be done either by disabling a default rule, or by intensifying a default rule (by moving it upwards in the ordering relation), or by attenuating a default rule (by moving it downwards in the ordering relation). We can exemplify this with an exclusionary reason, i.e. a higher-order reason that says we should ignore another reason. Horty uses 'Out(d)' to stand for the operation of excluding a default rule (Horty, 2012, p. 125). So, an exclusionary reason is such that it triggers a default rule of the sort 'X $\rightarrow$ out(d_n)'.

Take the following example: Let D stand for 'Hamilton promised to have lunch with Burr', E stand for 'Hamilton promised to have lunch with Adams', and F stand for 'Hamilton's promise to have lunch with Burr was given under duress'. Let G stand for 'Hamilton has lunch with Burr'. Now let the d_3 be the default rule 'D $\rightarrow$ G', d_4 stand for 'E $\rightarrow$ not-G', and d_5 stand for 'F $\rightarrow$ Out(d_4)'.

Let us assume the following ordering relation '$d_3 < d_4 < d_5$' and that D, E, and F all obtain.

We need to note that Horty (2012, p. 125) complicates the triggering condition and assumes that a default rule is triggered just as long as the premise of the rule obtains and the rule is not excluded.

Given that $d_3 < d_4$, Hamilton's reason E to have lunch with Adams would be stronger than Hamilton's reason D to have lunch with Burr. However, since F also obtains, the default rule d_5 is also triggered, and therefore d_4 is excluded. Given our example and our modified triggering conditions, only d_3 and d_5 are triggered, so Hamilton ought to have lunch with Adams.

Horty's view has the advantage of being theoretically austere, and of side-stepping most of the problematic aspects of the weighing metaphor.

The pressing question about Horty's account is this: How are we to determine the ordering relation among default rules? The weighing metaphor may be inexact, but we are at least given some kind of story about ordering – we have an intuitive grasp of what it means for a reason to be weighty. On Horty's framework, by contrast, we are given a working framework but no real clue as to how to determine the ordering relations of the default rules involved. On the other hand, perhaps this is as it should be. Perhaps all Horty intends to offer us is a framework

which needs to be supplemented with other normative theories. It is my hunch, but for the time being mere speculation, that it is possible to unify Horty's framework with a more standard conception of the weight of reasons.

The view that Horty account is primarily a framework is satisfyingly compatible with my claim that debates over how to understand the weighing metaphor for reasons are more often than not about arriving at overall reasons, rather than weighing reasons. The final weighing of overall reasons is seen as a quite simple and unproblematic process: Pair-wise comparison of the weight of the two overall reasons, or in Hortyian terms, an ordering of default rules. That said, determining the strength of the overall reasons, and arriving at an ordering of any default rules, may not be a simple task.

Having explored what might be called the internal workings of reasons, and having asked what reasons there are and how to weigh them, in the next chapter we will look at the way reasons relate to what we ought to do and what is valuable.

On Buck-Passing

Some of the more important and well discussed issues in contemporary normativity research focus on questions about how reasons relate to what is valuable, what we ought and are obligated to do, what it is rational or fitting to do, believe or feel, and so forth.[54] A vital part of these discussions is the question of whether some normative phenomena are more fundamental than others. That is, whether we can pass the buck from one normative phenomenon to another.[55]

Recently, numerous positions have been brought forth in the search for the fundamental normative phenomenon. We have monist positions such as *Reasons First*, which treat reasons as fundamental (Schroeder, 2007; Skorupski, 2010a; Parfit, 2011a; Scanlon, 2014). The *Value First* tradition treats value as the sole fundamental normative phenomenon (Moore, 1993/1903; Maguire, 2016). Additionally, Broome has recently argued for the position that what we ought to do is more fundamental than what we have reasons to do (Broome, 2013). [56] There are also the *Fittingness Firsters*, who hold that fittingness is more fundamental than reasons (McHugh & Way, 2016; Rowland, 2017). Finally, there are *pluralists* who allow for more than one fundamental normative phenomenon, and some of whom are hopeful that we can understand *some*

---

[54] While space does not allow me to discuss the relation between reasons and rationality, this relation centers on two issues: (1) Is rationality normative in the sense that something being rational entails that we have reasons to do it (see (Broome, 2005; Kolodny, 2005; Broome, 2007a, 2007b, 2013; Sylvan, 2015)); and (2) what reasons matter in rationality: the reasons we actually have, the reasons we believe we have, or the beliefs that we have which, if true, would be reasons (see (Scanlon, 1998, pp. 30-32; Broome, 2007a; Parfit, 2011a, pp. 111-125; Sylvan, 2015)).

[55] The phrase "pass the buck" was popularized in Scanlon (1998, p. 98).

[56] It might be a bit hasty to describe Broome's position as 'Ought First'. He does believe ought to be more fundamental than reasons, but to my knowledge he has never explicitly argued that 'ought' is the only fundamental normative phenomenon.

normative phenomena in terms of others but also think we need more than one fundamental normative phenomenon (c.f. (Dancy, 2004; Reisner, 2015b)).

In this chapter, I will concentrate almost exclusively on the Reasons First approach. I will first discuss whether we should pass the evaluative buck, i.e. understand value in terms of evaluative reasons. Then I will discuss whether we should pass the deontic buck, i.e. understand what we ought to do and the deontic status of an action – e.g. its being obligatory or permissible – in terms of practical reasons. The motivation for this delineation of scope is first and foremost that in the papers following this introduction I write primarily within the Reasons First framework. Secondly, space does not allow me to give a proper characterization of the intricacies of the other positions. The exception to my focus on a Reasons First approach is the discussion I present of whether we should understand what we ought to do in terms of what we have reasons to do. There I discuss Broome's proposal that what we ought to do is more fundamental than what we have reasons to do.

Evaluative Buck-Passing

In recent times one of the more popular ways to pass the buck is to understand an object's being (positively) valuable in terms of its being fitting to favor, or hold other pro-attitudes towards, that object.[57] This family of analysis is called the 'Fitting Attitudes Analysis of Value' (henceforth the FA analysis).[58] The FA analysis in its generic form can be formulated as follows:

---

[57] I here use "object" in a slightly idiosyncratic way to refer to anything that could be a bearer of value, i.e. that which has value. There are several views on which ontological categories can function as bearers of value – concrete objects, events, state of affairs, etc. (c.f. (Anderson, 1993; Rabinowicz & Rønnow-Rasmussen, 2000; Zimmerman, 2001)).

[58] The FA analysis has a long history going back to at least Franz Brentano, who understood an object being valuable in terms of love of that object being correct. Other historically important proponents of the FA analysis include (Broad, 1930; Ewing, 1948; Sidgwick, 1981/1874; Chisholm, 1986; Falk, 1986; Wiggins, 1987; Anderson, 1993). The FA analysis has seen renewed attention since Scanlon (1998). For some contemporary defenders of the FA analysis, see (Skorupski, 2010a; Parfit, 2011a; Rønnow-Rasmussen, 2011; Way, 2012; Garcia, 2018a).

> *FA analysis (generic):* For it to be the case that *x* is valuable is for it to be the case that it is fitting to (dis-)favor *x*.

It should be noted that there are two rather different claims that we can be making when we pass the buck. One can make a conceptual claim, such as "For any object *x*, what it *means* for *x* to be valuable is for *x* to be fitting to favor". Alternatively, one can make a metaphysical claim, such as "What it *is* for an object to have the property of being valuable is for the object to have the property that it is fitting to favor it". One can of course pass both the conceptual and the metaphysical buck, but one need not do so. In the debate, it is sometimes not made explicit whether the topic is the conceptual or metaphysical claim, which makes exegesis all the more taxing. In this chapter, my interest lies with metaphysical questions, such as whether we can give a reductive metaphysical explanation of *x* having certain axiological properties (such as being valuable) in terms of *x* having the property of being fitting to favor.[59]

Here 'favoring' and 'disfavoring' should be understood as placeholders for all pro-attitudes and con-attitudes – for admiring, loving, loathing, disliking, and so forth. Exactly what makes an attitude a pro or cons, and which of these attitudes are needed, are topics requiring more research.[60] Some value pluralists want to match each specific attitude with a specific value – e.g. if an object is admirable, the appropriate attitude is admiration. There seem to be at least four options here: i) a one-to-one match between the type of attitude and its corresponding value; ii) a one-to-many match between a type of attitude which it is fitting to have towards a plurality of values; iii) a many-to-one match between a plurality of attitudes which it is fitting to have towards a single value; and iv) a many-to-many match between a plurality of attitudes and a plurality of values.[61] I will henceforth

---

For more on the history of the FA analysis of value, see (Rabinowicz & Rønnow-Rasmussen, 2004; Danielsson & Olson, 2007; Gertken & Kiesewetter, 2017).

[59] In footnotes I will, however, note a few conceptual issues so as to hint at the broader issues raised by buck-passing.

[60] For discussion of this issue, see (Rabinowicz & Rønnow-Rasmussen, 2004; Garcia, 2018b).

[61] Not much has been written on how to match attitudes and values. (Mulligan, 1998) gives a good overview of the problem.

refer to the favoring and disfavoring part of the FA analysis (i.e. the pro- and con-attitudes) as the *attitudinal component* of the analysis.

The *normative component* of the FA analysis – the 'fittingness', in other words – can be given several distinct interpretations. The most common is in terms of reasons (Scanlon, 1998; Rønnow-Rasmussen, 2011), but other interpretations are available, some of which cast in terms of an ought, or a requirement (Ewing, 1948; Rabinowicz, 2008, 2012), and some of which treat fittingness as normatively fundamental (McHugh & Way, 2016). The entities that make it fitting to favor an object are understood, in the FA analysis, as value-makers, i.e. as what makes the object in question valuable. In Paper 5, Henrik Andersson and I argue that a plausible way of understanding the normative component is in terms of *conclusive reason*. This position is a development in the tradition of buck-passers who understand fittingness in terms of reasons, such as (Scanlon, 1998; Skorupski, 2010a; Rønnow-Rasmussen, 2011).

The FA analysis shows a lot of promise. If it is correct, it has the explanatory power to answer the general question of what it *is* to be valuable by reducing the property of being valuable to the property of being something it is fitting to favor. It also manages to supply us with answers to more specific questions, such as what it is for an object to have different kinds of value – an example would be the proposal that we understand an object's *being admirable* in terms of its being fitting to admire the object. In other words, the type of value an object has can be determined by the attitudinal component of the FA analysis.

This analysis also sheds light, not just on what it *is* to be valuable, but also on *why*, or *in virtue of what*, an object is valuable. That it is the case that a certain object *x* is admirable is to be understood as it being the case that it is fitting to admire *x*. Given our interpretation of the normative component of the FA analysis in terms of reasons, we are able say in virtue of what it is fitting to admire *x*: It is fitting to admire object *x* in virtue of the reasons, i.e. the specific reason-facts, there are for admiring *x*. The reasons to admire *x* ground, or metaphysically explain, the fact that *x* is admirable.

Metaphysical Grounding

Before we continue our discussion of the merits of the FA analysis and the objections that it faces, we need to pause and explain what we mean by the technical notion of *metaphysical grounding*.

Several terms have been used to capture the notion of grounding. They include: 'in virtue of', 'if and only if, and because', 'as a result of which', 'metaphysically explains' and 'makes'. What they all have in common is that they try to point to a specific type of non-causal explanation. Perhaps the easiest way to distinguish this type of explanation is via examples. Here are two from Wygoda Cohen (2019, p. 1):

1) Norma's cheek hurts because Helen just slapped her.

2) Norma's cheek hurts because (or in virtue) of certain neurological activity occurring in her brain.

Both make use of "because", but they provide different kinds of explanation. In the first case we are presented with a causal explanation, where we are given the causes of Norma's cheek hurting. In the second case, however, we are presented with a non-causal explanation. In other words; we are given the grounds of Norma's cheek hurting (Wygoda Cohen, 2019, pp. 1-2).

Some of the more paradigmatic examples used to exemplify grounding are:

Singleton: Singleton Socrates is grounded in Socrates.[62]

Euthyphro: Do the gods love the pious because it is pious or is it pious because it is loved by the gods?

Conjunction: The conjunction "A & B" is true in virtue of [A, B].

A feature is normatively fundamental in the metaphysical understanding if and only if it is not grounded in anything else that is normative. One can, in a sense, think of this as trying to find the smallest normative building block.

---

[62] Singleton Socrates is the set that has Socrates as its only member.

Grounding is usually taken to be a relation with certain logical properties: It is asymmetric, transitive, and irreflexive. It also has certain structural properties: It is non-monotonic and hyperintensional.[63] By this I mean the following:

The grounding relation *R* is:

> *Asymmetric* if and only if for any two items *x* and *y* if *x* is related to *y* by *R* then *y* is not related to *x* by *R*. For example: If the reasons to favor *x* grounds the value of *x* then the value of *x* does not ground the reasons to favor *x*.

> *Transitive* if and only if for any items *x*, *y* and *z*, if *x* is related to y by *R,* and *y* is related to *z* by *R* then *x* is related to *z* by *R*. For example: If [the reasons to favor Marta] grounds the [value of Marta], and [Marta's value] grounds [Marta being deserving of an award], then [the reasons to favor Marta] ground [Marta being deserving of an award].

> *Irreflexive* if and only if for any item *x,* *x* is not related to itself by *R*. For example: The value of *x* does not ground the value of *x*.

> *Non-monotonic* if and only if for any item *x*, *y* and *z*, if *x* is related to *y* by *R* then it is possible that [*x* and *z*] are not related to *y* by *R*. For example: [the reasons to favor *x*] ground [the value of *x*], but it need not be that [the reasons to favor *x* and the cup of coffee on my desk] together ground [the value of *x*].[64]

> *Hyperintensional* if and only if, for any items *x* and *y,* if *x* necessarily obtains if and only if *y* obtains and *x* is related to *y* by *R* then 'y in virtue of x' is true but 'x in virtue of *x*' is false. Even though *x* and *y* are necessary co-extensional the substitution of necessary co-extensional entities is not guaranteed to preserve truth. To see this, compare "It is true that grass is green because grass is green" and "Grass

---

[63] The formal properties just listed are those standardly taken to be associated with grounding. However, few within the grounding debate subscribe to all of them, and in fact most, if not all, have been questioned. For discussion of the formal properties of grounding, see (Schaffer, 2009; Correia & Schnieder, 2012; Fine, 2012; Bliss, 2013; Raven, 2013; Rodrigues-Pereyra, 2015). For discussion of the way in which the grounding relation is non-monotonic, see (Green Werkmäster & Smid, MS).

[64] A helpful way to think about this is in terms of defeasible reasoning and Horty's (2012) framework of reasons from the previous chapter. If Tweety is a bird, we can (defeasibly) infer the conclusion that Tweety can fly. If, however, we find out that Tweety is a penguin, we are no longer permitted to infer that Tweety can fly, even though she is a bird.

is green because it is true that grass is green"(Nolan, 2014, p. 151).[65] An example of a hyperintensional relation is, arguably, the explaining-relation: If x (fully) explains y then it is not the case that y explains x.[66]

Hyperintensionality is, I believe, commonplace in our everyday lives, but it is nonetheless hard to grasp, so it is helpful to compare these two propositions to understand this aspect:

A) Necessarily, all bachelors are unmarried men if and only if water is H_2O.

B) Necessarily, x is water if and only if x is H_2O.

As may be evident, proposition *B*, in contrast to proposition *A*, has an explanatory direction; x is water in virtue of it being H_2O.[67] So, while proposition *A* is true, there is no explanatory connection between 'all bachelors are unmarried men' and 'water is H_2O'. I believe that, to many of us, a natural reading of *A* would be that *A* is false while *B* is true. Bracketing our introductory logic courses, we read *B* as saying something hyperintensional, as explaining what it is in virtue of which x is water. Likewise, it is not unnatural to read A as false, because it offers a defective explanation of why all bachelors are unmarried men. If it turns out that the FA analysis is correct, then the reasons to favor x will be what explains in x's being valuable.

It is also standard practice to distinguish between full grounding and partial grounding – e.g. [A] partially grounds [A & B] while [A] fully grounds [A or B].[68]

Historically, moral philosophers were mostly interested in *supervenience*, and in particular how the normative supervened on the natural (Hare, 1952). While

---

[65] Hyperintensionality was first introduced by Cresswell (1975); for more on the interesting connection between hyperintensional logic and hyperintensional metaphysics, see (Fine, 2012; Nolan, 2014; Schnieder, 2016)

[66] Note the similarities between metaphysical grounding and constitutive grounds in Paper II.

[67] The connection between grounding and metaphysical explanation is an issue of some controversy; some claim the connection is an explanation, while others claim that it merely backs a metaphysical explanation (see (Trogdon, 2013, Ms)).

[68] For more on the logic and structure of grounding, and on full and partial grounding, see (Correia & Schnieder, 2012; Fine, 2012; deRosset, 2013; Trogdon, 2013; Schaffer, 2016; Wygoda Cohen, 2019).

supervenience is important and gives us some insight into how normativity functions, it is does not say anything about which phenomenon is more fundamental. Supervenience is here understood as necessary co-variation, so that there can be no change in the supervenient feature without a difference in the supervenience base (Kim, 1993, 1994). The problem is that necessary co-variation does not give us enough. In order for A and B to supervene on each other it is enough that in every possible world where A exists, B exists. The co-variance relation is symmetric in the sense that if A co-varies with B then B co-varies with A, and therefore we do not get the hyperintensional feature we want, and need, in order to explain why A and B exist in all the same possible worlds. Of course, it may be the case that most moral philosophers who have made use of the concept of supervenience had a hyperintensional understanding of supervenience in mind (c.f. (Garcia, 2018a)). It is possible that the switch in terminology from 'supervenience' to 'grounding' is more terminological than substance.

The Fitting Attitudes Analysis of Value

With this short detour into the metaphysics of grounding complete, let us return to the FA analysis of value. The analysis has several benefits. First, there is strong intuitive support for the idea that there is a connection between values and attitudes – e.g. liking that which is good and disliking that which is bad. On the FA analysis, this connection is no mere cosmic coincidence; the analysis explains how the connection is supposed to be understood. Second, the FA analysis has the hallmark of a good scientific analysis in that it opens up further research. In recent years, it has been used not only to understand what it is to be valuable, but also to illuminate and bring unity to some of the more fine-grained evaluative properties, such as:

> Final Value: For it to be the case that x has final value is for it to be the case that it is fitting to favor x for its own sake (Zimmerman, 2001; Rabinowicz & Rønnow-Rasmussen, 2004).

> Instrumental Value: For it to be the case that x has instrumental value is for it to be the case that it is fitting to favor x for the sake of its effects (ibid).

> Personal Value: For it to be the case that x is good for P is for it to be the case that is fitting to favor x for P's sake (Rønnow-Rasmussen, 2011).

Attributive Value: For it to be the case that *x* is valuable as an F is for it to be the case that it is fitting to favor *x* as an F (Skorupski, 2010a).

Betterness: For it to be the case that *x* is better than *y* is for it to be the case that we are required to prefer *x* to *y* (Rabinowicz, 2008, 2012).

The FA analysis has an advantage, in that it is sufficiently neutral when it comes to first- and second-order questions. As Andrés Garcia writes:

Analyses of value tend to run the risk of committing themselves to positions that should intuitively be left open. On the surface, the fitting attitudes analysis seems to avoid this risk to an admirable degree. It says nothing about first order questions about what objects have value or why, nor does it entail anything about the semantics of value statements, the epistemology of value judgments, or the ontological status of evaluative properties and relation (Garcia, 2018a, p. 22).

A salient advantage of the FA analysis of value is that it promises to improve our ontological economy. On this analysis, value is not a fundamental normative property; it is something that can be understood in terms of reasons for pro- and con-attitudes. In this sense, value is no more mysterious than reasons are. This is not to suggest that reasons themselves are not ontologically cumbersome. It is just to observe that there is, at least, one less entity foisting additional mystery on to our ontology.[69]

---

[69] For more on the ontological problem of reasons, see (Parfit, 2011b; Olson, 2018).

A whole host of objections have been brought to bear on the FA analysis. Examples include the wrong kind of value problem,[70] the objection from polyadicity[71] and the objection from circularity. [72]

---

[70] The wrong kind of value problem is a family of objections all of which challenge the biconditional that an object is valuable if and only if it is fitting to favor it from left to right, so to speak. They argue that there are cases where something is valuable but it is nonetheless *not* fitting to favor it. One example is the solitary goods objection, which focuses on cases where there is something that we would intuitively claim is valuable, but there are no agents to favor it, so there are no agents for whom it is fitting to favor it, and hence it cannot be fitting to favor it. For more on this objection, see (Bykvist, 2009; Orsi, 2013). Another is the casual entanglement objection, which is that there are cases where some *x* would be good, and therefore according to the biconditional, fitting to favor, but the goodness of *x* would be undercut if anyone favored it. For more on this objection, see (Reisner, 2015a).

[71] The polyadicity problem notes that value is most often regarded as a monadic property while being fitting to favor is, at least, a two-placed relation, i.e. it is fitting *for an agent* to *hold some attitude* towards *x*. For more on this objection, see (Olson, 2013).

[72] The objection from circularity was originally presented by W. D. Ross (1963/1939) and has been frequently discussed since its introduction (Ewing, 1948, p. 166f; Wiggins, 1987, p. 189; Rabinowicz & Rønnow-Rasmussen, 2004). Ross succinctly presents the charge against the FA analysis of value in *Foundations of Ethics* (Ross, 1963/1939) as follows:

> The phrase 'worthy of admiration', it appears to me, does not justify the conclusion that the goodness which is so described is a relational property. For admiration is not a mere emotion; it is an emotion accompanied by the thought that that which is admired is good. And if we ask on what ground a thing is worthy of being thought to be good, only one answer is possible, namely that it is good. It would be absurd to say that a thing is good only in the sense that it is worthy of being thought to be good, for our definition of 'good' would then include the very word 'good' which we were seeking to define (Ross, 1963/1939, pp. 261 -262).

What Ross is saying is that 'admiration' cannot be understood without using the concept of 'good'. A minority of commentators have embraced the view that the FA analysis of value is circular but argued that not all circular analyses need be incorrect, uninformative, or philosophically uninteresting c.f., (Garcia, 2018a). Others, most prominently Ewing (1948) and (Rabinowicz & Rønnow-Rasmussen, 2004), have instead argued that the goodness of an object is not what makes the object fitting of admiration; rather it is the features that make the object good that make it fitting to admire it. In response to the circularity objection, it is argued that admiring an object need not be followed by a thought that the object is good (D'Arms & Jacobson, 2000). For instance, it seems possible to be a sceptic about goodness in general and still be able to feel admiration.

I do not intend to explore each and every objection. Some, such as the objection from circularity, are chiefly aimed at the conceptual thesis of the FA analysis, and some, such as the wrong kind of value problem, would require me to go into too much detail. However, I will discuss what has been regarded as the main challenge of the FA analysis, namely the 'wrong kind of reason problem' (henceforth the WKR problem).

The Wrong Kind of Reason Problem

The following biconditional is an implication of the FA analysis:

> *FA biconditional*: An object is valuable if and only if it is fitting to favor it.

The WKR objection is that the right side of the biconditional can be true without it being the case that the left side is true. Therefore, according to this objection, the FA analysis cannot be sound. [73] In order to put some flesh on these bones, consider the following example, paraphrased from Crisp (2000b, p. 459).

> *Saucer of mud*: There exists a saucer of mud. Initially you have no reasons to favor it, and therefore deem it as having no value. However, an evil demon threatens you with excruciating torture unless you favor the saucer of mud.

In *Saucer of mud* the demon's threat seems to give you a reason to favor the saucer of mud – it may even be a conclusive one. According to the FA analysis, the saucer of mud is therefore valuable. The problem is that intuitively the saucer of mud is without value, and even if it had value, the demon's threat would not make it more valuable. Hence one has a reason of the wrong kind to favor the saucer of mud.

---

[73] The WKR problem was first presented by Crisp (2000b), a similar worry is found in D'Arms & Jacobson (2000). The name of the objection and its paradigmatic formulation can be found in (Rabinowicz & Rønnow-Rasmussen, 2004).

The WKR problem can be formulated as the conjunction of four claims:

1. *X* has no value.

2. The demon's threat is a (conclusive) reason to favor *x*.

3. If there is a (conclusive) reason to favor *x*, then *x* is valuable.

4. *X* has value.[74]

Since 1 and 4 (which last follows from 2, 3) are in contradiction, proponents of the FA analysis need to deny or modify at least one of claims 1–3 in order to dissolve the contradiction.

A first maneuver might be to insist that while the saucer of mud is valuable, it is so only in an indirect way, as an object you can direct a favoring attitude towards in order not to be tortured. In other words, the saucer of mud has instrumental value. Rabinowicz and Rønnow-Rasmussen (2004, p. 403) has argued, however, that we easily set up a scenario in which we are explicitly dealing with final, rather than instrumental, value.

Say that the evil demon threatens you with excruciating torture unless you favor the saucer of mud *for its own sake.* Again, intuitively, we want to deny that the saucer of mud is finally valuable and at the same time say that you have a reason to favor the saucer of mud for its own sake.

In the last decade a plethora of different solutions to the WKR problem have been presented (Olson, 2004; Stratton-Lake, 2005; Danielsson & Olson, 2007; Lang, 2008; Skorupski, 2010a; Heuer, 2011; Parfit, 2011a; Way, 2012), and these have provoked an almost an equal number of papers attempting to show why none of the proposed solutions succeeds (Rabinowicz & Rønnow-Rasmussen, 2004, 2006; Olson, 2009; Reisner, 2009a; Schroeder, 2010). Space does not allow me

---

[74] Of course, it need not be a threat from a demon that is the wrong kind of reason. There are plenty of ways to generate reasons of the wrong kind. Nor is it necessary that *x* has no value. It can have a positive value or a negative value; what matters in cases where *x* has value is that the wrong kind of reason does not change the value. To illustrate, Derek Parfit is admirable, but a threat from my supervisor designed to get me to admire him is a reason of the wrong kind to admire Derek Parfit.

to explore all the intricacies of the debate, but I do have space to outline the three main strategies adopted to meet the WKR objection.

(i)   *Solving the problem*: Modify the FA analysis. Argue that while the threat is a reason to favor the saucer of mud, it is the wrong kind of reason, and insist that the FA-analysis only applies to the reasons of right kind (i.e. modify 2 and 3).

(ii)   *Dissolving the problem*: Argue that the threat does not give the threatenee a reason to favor the saucer of mud, so the objection never gets off the ground (i.e. deny 2),

(iii)   *Rejecting the FA analysis* (i.e. deny 3).

I take these in turn.

Solving the Problem

Those who want to solve the WKR problem propose to do so by distinguishing the reasons of the right kind from reasons of the wrong kind, and by arguing that it is only the former which matter when it comes to value. Explicitly including this modification, the FA analysis reads as follows.

The Fitting Attitudes Analysis RKR: For *x* to be valuable is for there to be (conclusive) reasons of the right kind to favor *x*.

There have been numerous attempts to draw the distinction between reasons of the right kind and reasons of the wrong kind. Thus, Schroeder (2010) holds that the right kind of reason is one that is necessarily shared by everyone engaged in the relevant activity. In the case of the saucer of mud, if someone who has not been threatened by the demon and is engaged in the activity of favoring does not have reasons to favor the saucer of mud, then the demon's threat is not the right kind of reason to favor the saucer of mud. Rabinowicz and Rønnow-Rasmussen (2004) instead propose that reasons of the right kind are those that are able to play a dual role – that of justifying the attitude and that of being suited to figure in the intentional content of the attitude. For example, Parfit's authorship of *Reasons and Persons* is a reason of the right kind to admire him because it justifies the admiration and can be treated as a fact (among others) for which I admire him.

The most orthodox way to draw the distinction between the right and the wrong kind of reason is by distinguishing state-given reasons from object-given reasons. To my knowledge, this distinction was first proposed by Parfit (2001), although he used it for a purpose other than to solving the WKR problem. In this approach, state-given reasons are reasons given by certain features of being in a certain state – say, that of favoring the saucer of mud. Object-given reasons are reasons given by certain features of the object of the attitude. According to the state-given/object-given solution to the WKR problem, all reasons of the wrong kind are state-given reasons. It is the object-given reasons that are reasons of the right kind.

The proposal is *prima facie* plausible, because reasons of the wrong kind seem to be those that matter for the value of *holding* the attitude in question and otherwise spring from costs and benefits of holding the attitude in question.[75] At the same time, reasons of the right kind seem to be those that matter for the value of the object. However, as Olson (2004, p. 299) points out:

> [I]t is easily seen that properties of the attitudes may be recast as properties of the objects: if, e.g., the attitude of preferring the saucer of mud has the property of preventing our suffering of severe pain, then the saucer of mud has the corresponding property of being such that prefacing it would prevent out suffering severe pain.

Furthermore, Schroeder (2012) has described cases where it is plausible to maintain that there are state-given reasons that are nonetheless reasons of the right kind.

Although all these attempts show some promise, all of them, and indeed others in the same vein (e.g. (Stratton-Lake, 2005; Lang, 2008)), have drawbacks.[76] So while it might be possible to give adequate non-*ad hoc* demarcation criteria that separate reasons of the wrong and the right kind, we can agree with Gertken and

---

[75] A classic example of this is Kavka's (1983) Toxin Puzzle. Roughly speaking, in this puzzle we seem, intuitively, to have a reason to be in a state of intending to drink poison in order to win a prize but have no reason to carry out the intention, i.e. drink poison, because all that is needed to win the prize is to have the intention.

[76] For a good overview of these attempts, see (Rabinowicz & Rønnow-Rasmussen, 2004, 2006; Gertken & Kiesewetter, 2017).

Kiesewetter that "it is safe to say that none has so far gained wide acceptance" (2017, p. 4).

Dissolving the Problem

Proponents of dissolving the problem (e.g. (Skorupski, 2010a; Parfit, 2011a; Way, 2012)) are usually called 'WKR skeptics'. They deny that there are any reasons of the wrong kind – thus, in response to *Saucer of mud* they deny that you have any (conclusive) reason to favor the saucer of mud. You only have a (conclusive) reason *to get yourself* to favor the saucer of mud. In other words, you have a practical reason to perform a certain act, i.e. the act of getting yourself to admire the saucer of mud. The WKR skeptics argue that the FA analysis is only concerned with evaluative reasons to have certain pro- and con-attitudes. This response is explicit in Skorupski's discussion of a case of admiring a terrible violin performance:

> [T]the fact that the evil demon has evil plans is a sufficient reason for me to *do* something – namely, bring it about that I admire the performance, if I can. In the circumstances, that would be a very good thing to bring about. Over and above that uncontroversial point, there is then the question of whether an evaluative reason relation also holds. Does the selfsame fact about evil demon stand in that *distinct* reason relation to me and a certain feeling of mine, namely, *admiring the performance*? The two relations are distinct, since their *relata* are distinct. And once they have been distinguished, a case needs to be made for holding that the second relation holds as well as the first (Skorupski, 2010a, p. 88).

While the WKR skeptic's solution is simple and elegant, it faces a simple problem. Intuitively, the evil demon's evil plan is a reason for us to admire the performance. The WKR skeptic has the burden of explaining why we are mistaken about this. Even if you have a reason to bring it about that you admire the performance, it seems that a simple explanation for this reason is the fact that you have a reason to have the attitude. If you manage to admire the demon, the WKR skeptic is forced to say that you have an attitude that you have no reason to have.

I believe that the WKR skeptic would explain our allegedly mistaken intuitions in mud-like cases by agreeing that in *most* cases the explanation of why we have a reason to get ourselves to have the relevant attitude is that we have reasons to have that attitude, but adding that this is not always the case. In fact, (the WKR

skeptic) Skorupski has proposed a connection between evaluative reasons and practical reasons for actions. He calls it a "bridge-principle", and it is as follows:

> *The Bridge-Principle*: If an agent, A, has a reason to have attitude *x* then, and because of that, A has a reason to do *x*-prompted actions (Skorupski, 2010b, p. 129).[77]

According to this principle, the reason for A to have a certain attitude transmits to the attitude-prompted actions.[78] Skorupski does not give an exhaustive account of what an attitude-prompted action is, but he describes such actions as the acts that one is "characteristically disposed" (Skorupski, 2010b, p. 129) to perform when one has a certain attitude. He gives a few examples: "Out of sheer goodness of heart, someone does me an unrequested good turn. That fact certainly gives me a reason to feel grateful. And because I have reasons to feel grateful to him for his good turn, I have reason to act from that gratitude, for example by thanking him or giving him a present or by returning the favour" (Skorupski, 2010b, pp. 129 - 130).

Presumably, "getting yourself to have the attitude" that you have reasons to have would qualify as an attitude-prompted action. However, the bridge-principle only goes in one direction – from evaluative reasons we can derive practical reasons, not the other way around. So, while the most common way to acquire reasons to get yourself to admire a performance is by having reasons to admire the performance, it is nonetheless possible to have reasons to get yourself to have an attitude that is not grounded in a reason to have that attitude. In the evil demon case, the reason to bring it about that you admire the performance is that otherwise you will be tortured.

---

[77] I adopt, and interpret, Skorupski's usage of "because of" as indicating that the evaluative reasons serve as the ground for the practical reason.

[78] I believe there are quite a few ways in which having a reason to do (universal verb) something transmits to other actions (universal verb). In Paper 4, "Normative Transmission", I discuss whether having a reason to do a certain action gives one a reason to take the necessary means to do that action. In addition to transmission principles concerning 'reasons for attitudes and attitude-prompted actions' and 'ends and necessary means', it is plausible that there also are principles concerning 'ends and sufficient means' (c.f. Kolodny (2018)), 'ends and optimal means', 'prudential reasons and moral reasons' and potentially a whole host of other principles.

By analogy, consider the theoretical domain. Most would agree that a demon threatening me with excruciating torture unless I believe p is not a reason to believe p – it is not a theoretical reason to believe p, but a practical reason to get myself to believe p.[79] Why would the evaluative case be any harder than the theoretical?

Rejecting the FA Analysis

So far we have discussed solving, and then dissolving, the WKR problem. Another solution is, of course, to reject the FA analysis understood in terms of reasons, i.e. to deny the claim made in the third premise that if there is (conclusive) reason to favor an object, then the object is valuable. Recently, largely in response to the WKR problem, some commentators have indeed rejected the versions of FA analysis of value where the normative component is formulated in terms of reasons. Instead, they have proposed that we should understand fittingness as normatively fundamental (McHugh & Way, 2016). This makes it possible to accept that the demon's threat is a reason to favor the saucer of mud, but still reject the notion that it is fitting to favor the saucer of mud. On their view, reasons do not make attitudes fitting. The demon's threat is a reason to admire the saucer of mud, but it is not something that makes it fitting to admire the saucer of mud. Alternatively, one could, as (Maguire, 2016, 2017) has proposed, understand value as more fundamental than reasons, so that we understand reasons to favor in terms of what is valuable. While these positions are in no way straightforwardly implausible space does not allow me to evaluate them properly in this

---

[79] An exception to this is, of course, presented by pragmatists such as (Reisner, 2009b, 2018a); but it is unclear if those who believe that, say, Pascal really does have a (pragmatic) reason to believe in the existence of God would claim that pragmatic reasons are of the wrong kind. One believes that pragmatic reasons for beliefs are reasons of the wrong kind only if one believes (i) that there are pragmatic reasons for beliefs, and (ii) that it is never the case that you ought to believe something based on pragmatic reasons for belief. If one takes the view that sometimes we ought to believe something in virtue of pragmatic reasons for beliefs, the view is more akin to biting the bullet – claiming that the saucer of mud is in fact valuable rather than trying to solve the WKR problem – i.e. unless one further believes that the right kinds of reason for beliefs are reasons that concern the truth of a propositions and the wrong kinds of reason are reasons that do not. To add "truth", however, would be equivalent to claiming that the right kinds of reason are those that concern the value of x and the wrong kinds of reason are those that do not – in which case the FA analysis would be seriously circular.

introduction. Since they have a minor role, if any, in the following papers, I will not discuss them further.

Summary

The debate surrounding the WKR problem has been running for more than fifteen years, and so far no consensus on the issues has been reached.

As may be evident, there are, as yet, no perfect solutions to the WKR problem. However, given the virtues and explanatory power of the FA analysis, it seems that rejection of the analysis would be ill advised. I believe that the present indications are that dissolving the problem is most likely to emerge as the correct strategy. It is not for me to settle this debate over the FA analysis here. It is, however, my hope and my aspiration that in the future I will be able to contribute to the burgeoning literature on the topic and shed new light on the WKR problem.

Deontic Buck-Passing

I have so far looked at attempts to pass the evaluative buck, and to understand evaluative properties in terms of deontic ones. An at least as interesting endeavor is to pass the deontic buck. Buck-passing of this sort has not been as intensively discussed as its evaluative counterpart.

This section has five parts. In the first, I first discuss understanding what we ought to do in terms of what we have (practical) reasons to do.[80] In the second, I propose an attempt to extend the analysis, cast in terms of reasons for actions, to the deontic properties of *being obligatory*, *being permissible*, *being impermissible* and *being optional*. In the third part, I discuss a few objections to the analysis presented in the previous part and propose various modifications to the analysis that may answer these objections. In the fourth part, I discuss an objection according to which we cannot understand what we ought to do in terms of what we have reasons to do because the 'can' in the dictum 'ought implies can' is different from the 'can' in 'reasons imply can'. In the end, I argue, the objection fails. The fifth

---

[80] Note that in this section I am using "do" not as a universal verb, but strictly to refer to actions.

part of this section summarizes the conclusions I have drawn and paves the way for the final chapter of the extended introduction.

As was the case in the previous section on evaluative buck-passing, my main focus will be on positive aspects of the buck-passing account. I look primarily at how one can pass the deontic buck, rather than focusing on negative questions, such as why other accounts, like Broome's *ought fundamentalism* are unsatisfactory (see page 27).[81]

The Standard Account

I take the following entailment to be almost universally accepted.[82]

> *The Minimal Thesis*: If an agent A ought to φ then A has a *pro tanto* reason to φ.

If we can pass the deontic buck, we can explain the Minimal Thesis.[83] That is, we ought to be able to explain the Minimal Thesis by appealing to the following general characterization of the connection between ought and reasons:

> *The Standard Account*: For it to be the case that an agent A ought to φ is for it to be the case that A has a conclusive reason to φ.[84]

---

[81] There are a few other lines of thought as well that deserve a more thorough treatment than I am able to give in this thesis. For example, (Kearns & Star, 2009, 2013) wish to understand reasons in terms of evidence of oughts, and others have proposed that we should understand reasons as premises of good reasoning (Setiya, 2014; Way, 2017).

[82] This conditional is near-universally accepted; however, (Horty, 2012, pp. 41 - 47) has argued that there are some special cases where it fails.

[83] Although the Minimal Thesis is almost universally accepted, Toni Rønnow-Rasmussen, in personal communication, has pointed out that it seems to violate conversational maxims. If I am to guess which conversational maxim is being violated, the most likely culprit is the *maxim of relevance* (Grice, 1989, pp. 26 - 27). While the Minimal Thesis is true, asserting it would elicit confusion and a response along the following lines: Sure, the fact that an agent A ought to φ implies that A has a *pro tanto* reason to φ, but it is not *just a pro tanto* reason but also a conclusive reason.

84 Versions of the Standard Account are frequent in the literature; for some prominent examples, see (Schroeder, 2007; Bedke, 2011, p. 130; Parfit, 2011a, p. 33). To remind the reader, in on page 28 of this thesis the notion of a conclusive reason to φ was defined as the reason (*pro tanto*

If you accept the Standard Account you pass the deontic buck from ought to reason. In the chapter "The Reason-Relation", we concluded that having a conclusive reason to φ involves having a *pro tanto* reason to φ. So, the Standard Account can explain why the Minimal Thesis holds. It holds because it is a logical consequence of the Standard Account.[85]

In order to explain fully what is involved in passing the buck from oughts to reasons, I need to say a few words about what senses there are of 'ought', and about what specific deontic property I am interested in. 'Ought' can refer to a multitude of specific properties, some of which relate merely to norms and convention, and others of which relate to specific domains (i.e. to what we ought to do, believe, or feel) or to specific normative spheres like the spheres of morality and prudence.

Take the following statements:

> *Chess*: "You ought not to move the F-pawn."

> *Promise*: "Morally, you ought to keep your promise."

> *Review*: "Yes, (morally) you ought to review the paper, and (prudentially) you ought to take a day off – but what you *really* ought to do in the end is to review the paper."

It is unclear whether 'ought' refers to the same deontic property in all three of these examples. In *Chess,* the ought appears to be connected to norms and conventions governing the board game. In *Promise,* the ought appears to be an explicitly moral property of a certain action. In *Review,* the ought appears to be being used both to designate what you ought to do relative to certain spheres of normativity, and what you ought to do considering all the spheres pertaining to

---

or overall) which outweighs all other reasons (*pro tanto* or overall). The conclusiveness says something about the strength of the reasons to φ as compared with the strength of the reasons to not-φ.

[85] Note that this is a buck-passing account of ought in terms of practical reasons. It is not, however, a fitting-attitudes account of ought in terms of reasons for the simple reason that no attitudes, nor any evaluative reasons, are mentioned. Since "buck-passing account" and "fitting-attitudes account" are sometimes used interchangeably, this clarification is worth highlighting. I thank Toni Rønnow-Rasmussen for pointing this out.

the domain of practical normativity. This ought is what Broome (2013) calls the *central ought*. [86]

In this section, we are concerned primarily with Broome's central ought, and more specifically with the central *practical* ought, i.e. what we ought to *do*. This is for two main reasons. First, mere norm- and convention-guided properties might behave differently from the central ought; and secondly, sphere-specific oughts (e.g. moral or prudential oughts) might behave differently depending on the particular normative sphere. My focus on the practical ought, rather than other normative domains such as the evaluative ought or the theoretical ought, is adopted simply because most of the literature that I will be discussing is on this ought. To cover the oughts of other domains, and in particular the theoretical domain, I would need more space than this thesis allows. Some aspects of the oughts associated with the theoretical domain and the evaluative domain will come up in the next chapter.

Bernard Williams described the central ought (using the term 'deliberative ought') as the end of deliberation, the ought you reach when there is no further question to ask as regards what you ought to do (1981b). We can imagine someone saying that, morally, she ought to keep her promise. We can imagine that same person going on to say that, prudentially, she ought not to keep her promise. We can also imagine the person ending up asserting that, when everything is said and done, the thing she ought to do is to keep her promise. It is this all-things-considered normative property that I am primarily concerned with in this section. The moral ought, however, will also play a prominent role when we turn to the connection between what we ought to do and what we are obliged to do.

So, according to the Standard Account we can pass the buck from ought to reasons. We will now discuss the possibility of extending the Standard Account to other deontic properties.

---

[86] Other terms that have been used to denote the (for all intents and purposes) same property are "final ought" (Reisner, 2018b) "all-in ought" (Bedke, 2011) and "deliberative ought" (Williams, 1981b).

The Properties of being Required, Permissible, Impermissible and Optional

The question is, then: Can we extend the Standard Account in such a way that it allows us to understand other deontic properties such as being obligatory, being permissible, being impermissible and being optional in terms of reasons?

As said previously, the Standard Account is as follows:

> *The Standard Account*: For it to be the case that an agent *A* ought to *φ* is for it to be the case that *A* has a conclusive reason to *φ*.

The reasons referred to in the Standard Account are *A*'s practical reasons, i.e. reasons for acting. So, the practical ought, the ought relating to the practical domain, is understood in terms of reasons for actions. The deontic properties we want to understand now are *obligatory, permissible, impermissible*, and *optional*. I will use the following abbreviations: *Obligatory (OB), Permissible (PE), Impermissible (IM),* and *Optional (OP)*.

The following assumption seems to be commonly made, both in our common-sense thinking and within philosophy:

> *Ought-Obligation (O-OB)*: For it to be the case that *A*'s *φ*ing is obligatory is for it to be the case that *A* ought to *φ*.

If O-OB is true, passing the deontic buck from ought to reasons means that we get a buck-passing account of obligatoriness for free. That is not all, however. According to standard deontic logic, the deontic properties we are interested in have the following implications.

> An act *φ* is:
>
> *Permissible* if and only if it is not obligatory to not *φ*, i.e. PE(*φ*) if and only not-OB (not-*φ*)
>
> *Impermissible* if and only if it is obligatory to not *φ*, i.e. IM(*φ*) if and only if OB (not-*φ*).
>
> *Optional* if and only if it is not obligatory to not *φ* and not obligatory to *φ*, i.e. OP(*φ*) if and only if (not-OB(*φ*) and not-OB (not-*φ*)).

Although it is possible to reject these logical relations, especially considering the discussion to come, they are widely accepted (McNamara, 2014).

As we can see straight away, it is possible to understand permissibility, impermissibility and optionality in terms of obligatoriness. So, if we accept O-OB and the logical relations between the deontic properties we are interested in, not only do we get to a buck-passing account of obligations in terms of reasons, we also get a buck-passing account of all the properties we were interested in. That account looks like this:

> *Obligation*: For it to be the case that A's φing is obligatory is for it to be the case that A has a conclusive reason to φ.

> *Permissibility*: For it to be the case that A's φing is permissible is for it to be the case that A lacks a conclusive reason to not-φ.

> *Impermissibility*: For it to be the case that A's φing is impermissible is for it to be the case that A has a conclusive reason to not-φ.

> *Optionality*: For it to be the case that A's φing is optional is for it to be the case that A lacks a conclusive reason to φ and lacks a conclusive reason to not-φ. [87]

While neat, passing the deontic buck this way runs into a few objections. For instance, there seem to be nuanced constructions in ordinary language where we distinguish between what we ought to do and what we are obligated to do. We can, for instance, imagine the following sign on the door to a restaurant bathroom:

(1) Employees must wash their hands.

(2) Customers ought to wash their hands.

If we interpret the 'must' in (1) as outlining an obligation for employees to wash their hands, then if *O-OB* is correct we get the result that both (1) and (2) refer to the same deontic property. A more charitable interpretation is that either (1)

---

[87] In the first chapter we noted that having a sufficient reason to φ implies that you lack a conclusive reason to not-φ and vice versa. For this reason it is possible to rephrase the definition in terms of sufficient reasons. On such a rephrasing permissibility, for instance, would come out as it being the case that you have a sufficient reason to φ.

and (2) refer to different properties or there is a difference in implicature between (1) and (2). The details of the specific example need not bother us now. All that the example is meant to show is that we cannot accept O-OB and regard our work as done. Since O-OB violates some part of our common language usage, there is a further burden of explanation on its proponents.

In what follows, I will present a few objections to O-OB. I believe that all of the following objections to our complete buck-passing account of the deontic properties OB, IM, PE, and OP stem from acceptance of some form of the following principle:

> *Obligation-Blame* (*OB-Blame*): If A is obligated to φ and A does not φ then A is blameworthy for not φing.[88]

The objections have the structure of a *reduction ad absurdum*. They start by (1) identifying a case where we have a conclusive reason to φ. Then (2) the Standard Account is invoked to draw the conclusion that we ought to φ. Then (3) by invoking O-OB we get the conclusion that we are obligated φ. The conclusion that we are obligated to φ, however, seems too heavy-handed, because (4) by invoking *OB-Blame* we see that we are now blameworthy for not φ-ing. And given the case that was identified in (1), this looks wrong – we do not want to claim that we are blameworthy for not φ-ing.

While I believe that there are possible replies to these objections, I will conclude, in the end, that the lesson to be learned from them is that *O-OB* is false (but that the Standard Account and *OB-Blame* are true).

---

[88] I want to stress that OB-Blame as it stands is severely simplified. Plausibly, a proper account of the connection between blame and obligation will involve some sort of epistemic condition, and perhaps (more controversially) a clause outlining extenuating circumstances, i.e. excuses. I am not opposed to either of these. The reason that I have not included them is because these conditions do not become relevant in the discussion to follow. The virtue of presenting OB-Blame in a succinct manner, I think, outweighs its drawbacks. I believe that formulating it this way makes the line of reasoning to come easier to follow. For discussion of the epistemic conditions of blame, and extenuating circumstances in general, see (Strawson, 1962; Wallace, 1994; Fischer & Ravizza, 1998; Shoemaker, 2011; McKenna, 2012; Mason, 2015; Rossi, 2016).

Enticing Reasons and Supererogatory acts

If both *O-OB* and *OB-Blame* are accepted, it follows that what we ought to do is what we are obligated to do, and what we ought to do is also what we are blameworthy for not doing.[89]

There seem to be two classes of case where it is usually assumed that the balance of reasons favors φing but it is permissible to not-φ and you are not blameworthy for not-φing. We need to keep in mind that this entails that we need to reject either the Standard Account, or *O-OB*, or *OB-Blame*. The two classes of case I have in mind, exemplified below, are cases of *supererogation* and cases of *enticing reasons*.

> *Ice-cream*: You can pick either chocolate ice-cream or strawberry ice-cream. You have a conclusive reason to pick chocolate ice-cream. However, given the trivial nature of the choice it is permissible for you to pick strawberry (Bedke, 2011, p. 130).

> *Burning orphanage*: An orphanage is burning. You could go in and try your best to save the ten children that are trapped inside the blazing inferno. Doing so would put you at a great risk of being killed. You have a conclusive reason to try your best to save the ten children. The personal risk, however, is so great that it is beyond the call of duty to require this of you. It is therefore permissible for you not to try your best to save the children. It is also very praiseworthy of you if you do it (Urmson, 1958/1969; Bedke, 2011).

> *Trash*: Your neighbor forgot to put the garbage can on the curb, and has just gone on vacation. The balance of reasons favors your putting your neighbor's garbage can on the curb. Given that it was your neighbor's responsibility to do so, it is permissible for you not to do it. It is, however, both permissible and praiseworthy for you to do it.

Ice-cream is supposed to illustrate a case of enticing reasons; reasons so trivial that even if the balance of reasons favors one option it is not the case that you have an obligation, or ought, to do it, nor is it the case that you are in any way

---

[89] I take it for granted here that being obligated to φ is equivalent to being obligated to not not-φ, which entails that not-φing is impermissible and that φing is permissible. These are orthodox entailment relations within standard deontic logic. I wish to remain neutral on the question whether we should accept all the axioms of that logic.

blameworthy if you chose the other option. Enticing reasons, according to Dancy, are to be contrasted with peremptory reasons. The distinction between enticing and peremptory reasons can be found in (Raz, 1986; Dancy, 2004; Robertson, 2008).[90] According to these theorists, if you do not do what you have conclusive reason to do, but the reasons involved are merely enticing, then you are not necessarily blameworthy for not doing what you have conclusive reason to do. At most, Dancy claims, someone could accuse you of being "silly" (Dancy, 2004, p. 21). If it is true that an agent can have conclusive enticing reasons to φ and no reasons not to φ without it being the case the agent ought to φ, then the Standard Account is false. Dancy believes that the Standard Account is indeed false, but he does not think this is a problem since, according to him, all we need to do is modify the Standard Account slightly: It must be specified, in other words, that what we ought to do is what we have conclusive *peremptory* reasons to do (Dancy, 2004, p. 21).[91]

I am doubtful about this use of the distinction between enticing reasons and peremptory reasons. The main difference between enticing and peremptory reasons seems to be that one type grounds an ought and the other does not. The cases used to argue for the existence of enticing reasons are more often than not cases involving slight enjoyment or discomfort – not too much enjoyment, not too much discomfort. When the enjoyment is great enough, the enticing reason 'transforms' into a peremptory reason. Secondly, the type of criticism that we expose ourselves to when we do not act in accordance with what we have conclusive reason to do, even if the case is quite trivial, is more serious than the, admittedly vaguely characterized, charge of "being silly". If we believe we have

---

[90] A different but similar distinction can be found in the writings of (Gert, 2007, 2012). Joshua Gert claims that reasons have two kinds of strength, justifying and requiring. A reason which, on Gert's view, only has justifying strength, and no requiring strength, would be what Dancy and Raz call an "enticing reason".

[91] Dancy seems to accept O-OB; at least, he accepts that if A ought to φ, not-φing is impermissible. He writes:

> One might allow that there is a weak sense of 'ought' in which one ought to choose the most enjoyable way of spending the afternoon, where no consideration other than pleasure are at issue. But in this weak sense of ought', it might be both that one ought overall to choose this and that one is permitted not to choose it. And this is hardly a sense of 'ought' at all (Dancy, 2004, p. 21fn).

conclusive reason to do something, and yet we do not do it, we seem to be slightly irrational. As Broome says:

> I recently advised a guest that he ought to try a mangosteen, on the grounds that mangosteens taste delicious. That they taste delicious would have to count for Dancy as an enticing reason. Nevertheless, I believe I spoke correctly. I did not think my guest was obliged to try a mangosteen, and I thought it would have been understandable for him not to do so. For one thing, he might not have believed me when I told him he ought to. Still, had he believed me, and had he still declined to try a mangosteen, I would have thought that a mild lapse of rationality on his part (Broome, 2013, pp. 60 - 61).

It is interesting to notice that when Broome says "I did not think my guest was obliged to try a mangosteen" he opens up a response that differs from Dancy's. Instead of modifying the Standard Account, that is, we can deny O-OB. In support of this interpretation we can note, quite interestingly also, that in Broome's 2004 paper there is an almost identical paragraph, except that he also writes "'obliged' is more heavyweight" (Broome, 2004a, p. 39).[92]

I think Broome's line of reasoning applies to *ice-cream* and any type of enticing case. It is the case that we ought to pick the ice-cream flavor that we have conclusive reason to choose, but it seems too heavyweight to claim that we are obliged to do so.

There is, however, something else that speaks for the intuition driving the introduction of enticing reasons. The intuition in question is that we would have no right to in any way intervene or blame the person who chooses the 'lesser' ice-cream or does not try the mangosteen. I think that one possible explanation of why it sounds too heavy-handed to say that we are obliged to try mangosteens, or pick the best ice-cream flavor, is that we often use 'obligation' as shorthand for 'moral obligation'. If blame is an essentially moral sentiment, and fits only in

---

[92] The full rumination from 2004 is: "I recently advised a guest that he ought to try a mangosteen, on the grounds that mangosteens taste delicious. That they taste delicious would have to count for Dancy as an enticing reason. Nevertheless I think I spoke correctly. I did not think my guest was obliged to try I mangosteen; 'obliged' is more heavyweight. I did think he ought to try one, but I simultaneously thought it would be permissible for him not to. Dancy generously points out that 'permissible' can be used in a way that makes these thoughts consistent" (Broome, 2004a, pp. 38 - 39).

moral cases, we can explain why the resulting ought, or obligation, in the trivial enticing cases sounds heavy-handed. The explanation is that no moral reasons pertain in the situation. It is understandable that we, as moral philosophers, care mostly about moral cases and would therefore hesitate to ascribe an obligation in an enticing case. The lesson to be learned from consideration of enticing reasons may just be that not all central oughts are moral oughts or have a moral element.

This latter response is, in effect, the rejection of OB-Blame in favor of the following principle:

> *MOB-Blame*: If A has a moral obligation to φ, and A does not φ, then A is blameworthy for not φing.

We will come back to *MOB-Blame* shortly, but first let us *examine Burning Orphanage* and *Trash*. The objection is that there is sometimes either a difference between what we have conclusive reason to do and what we ought to do, or a difference between what we ought to do and what we are obligated to do.

The notion of *supererogation* was first introduced in analytic philosophy by J. O. Urmson (1958/1969), and the precise necessary and sufficient conditions for cases of supererogation differ depending on who in the literature you are reading. For simplicity, I will go with a very broad understanding of supererogation. I will assume that a case qualifies as one of supererogation if it satisfies the first three of the following criteria plus one or more of the others:

(i)   There are two actions, φ and not-φing.

(ii) φing is praiseworthy.

(iii)   You have a conclusive reason to φ.

(iv)   It is not the case that you ought to φ.

(v)   It is not the case that you are obligated to φ.

(vi)   Not-φing is not blameworthy.

(vii) φing is morally better than not-φing.

Prima facie Burning Orphanage and *Trash* are good candidates to meet all seven criteria. Plausibly, you *do* have conclusive reason to save the orphans from being burned to death despite the risk to your life, and you *do* have conclusive reason to help out your neighbor even though it will be tiresome doing so. In order for the Standard Account to be true all that needs to be the case is that the combination of (iii) and (iv) can never be true. According to my intuitions, it is in no way strange to claim that it is true, in *Burning Orphanage* and *Trash*, that you ought to perform the action you have conclusive reason to do. I admit, however, that there is a theoretical cost if we deny the possibility of (iii) and (iv) being true at the same time. The cost becomes steeper if, in addition, we want to accept O-OB, thereby being forced to say that (i)–(v) cannot be true at the same time. Furthermore, denying that it is possible for (i)–(vi) to be met at the same time seems a step too far. I do not believe that you are *blameworthy* in *Burning Orphanage* if you do not risk your life, or that you deserve to be blamed in *Trash* if you do not help out your neighbor. We therefore need to deny either OB-Blame or O-OB.

There is plenty to say in favor of either option. Let us start with the acceptance of OB-Blame and denial of O-OB. Take the property of *being blameworthy*. Just as the property of *being admirable* is a value-property, it seems that the property *being blameworthy* is also value, a value-property. If we accept the FA analysis of value, it should be clear that we should therefore understand an agent being blameworthy in terms of evaluative reasons rather than practical reasons, i.e. in terms of reasons for attitudes rather than reasons for actions.

A Fitting-Attitudes Analysis of Obligations

Assuming we want to accept OB-Blame, this may lead us to a point where we want to understand the connection between obligation and blame in a more robust sense. We may wish to understand what it is to be obligated to do something in terms of what it is blameworthy not to do. Perhaps, to have an obligation to φ just is to be blameworthy if one does not φ, and to be blameworthy for that just is for it to be the case that we have (conclusive) reasons to blame an agent for not φ-ing. This would allow us to understand obligation (and the accompanying deontic properties) in terms of reasons to blame.

FA-Blame: For it to be the case that *A* is obligated to *φ* is for it to be the case that there is (conclusive) reason to blame *A* if *A* does not *φ*.[93]

FA-Blame lacks the finer details of their accounts, but shares obvious similarities with the positions taken by Stephen Darwall in *The Second-Person Standpoint* (2006, pp. 27-28) and John Skorupski (2010a, pp. 78, 291-300) in *The Domain of Reasons*.[94]

If we accept FA-blame and the Standard Account, we can explain why you ought to *φ* in *Burning Orphanage* but are neither obligated to do so nor blameworthy if you do not do so. What you ought to do is governed by your practical reasons, and what you are obligated to do is governed by the evaluative reasons there are to blame you if you do not do something.

The explanation of why you ought, but are not obligated, to *φ* in *Burning Orphanage* is this: You have conclusive (practical) reason to *φ*, but you are not blameworthy if you fail to *φ* and therefore you are not obligated to *φ*, because there is no (conclusive) reason to blame you for not *φ*-ing.

The acceptance of O-OB instead makes it hard to account for condition (vii) in our characterization of supererogatory actions. Condition (vii) claims that *φ*ing is morally better than not-*φ*ing. The combination of the Standard account, O-OB and MOB-Blame implies that where there is a conclusive moral reason to *φ* there is also a moral obligation to *φ*. If it is true that in *Burning Orphanage* you have a conclusive *moral* reason to *φ*, but it is also insisted that you are not blameworthy if you do not *φ*, then this speaks in favor of rejecting O-OB. So, according to the combined trio, you would have a moral obligation to *φ* and therefore be blameworthy unless you *φ*ied. If, however, we reject O-OB and distinguish between what one ought to do and what one is obligated to do, we can explain how, in the *Burning Orphanage* case, we are entitled to say *both* that it is morally better for you to *φ* *and* that you ought to *φ* without committing ourselves to the claim that not-*φ*ing is impermissible and blameworthy. A possible explanation of this is that the great personal cost in store for you is so grave that there simply is

---

[93] For clarity and simplicity, in the attitudinal component of this analysis I only mention blame. However, it is entirely possible that there is a range of what might be called (see below) 'holding-accountable attitudes' – e.g. attitudes to others' violations of non-moral obligations.

[94] See also (Gibbard, 1990, p. 42; Mill, 1998, p. 93) and (Ewing, 1939, pp. 2-3).

not a (conclusive) reason to blame you if you do not φ. We cannot reasonably demand that you φ.

Now I want to point out that several substantive moral theories, including various consequentialist theories, imply that there *is* an obligation to φ in *Burning Orphanage*, and moreover that failure to φ would be blameworthy.[95] I do not want to take a stance on the substantive question whether, in the *Burning Orphanage* case, there is such an obligation. But I do want to show that if we want to allow for the possibility supererogation, and not rule it out at a meta-ethical level, we cannot accept the trio of the Standard Account, O-OB, and MOB-Blame.

The rejection of O-OB might suggest that what we are obligated to do and what we ought to do become too disconnected. This may be especially worrying given that, according to FA-Blame, our obligations are determined by evaluative reasons and oughts are fixed by practical reasons. One plausible reply to this worry – one that I cannot explore in detail – is to say that there is still a connection between what we ought to do and what we are obligated to do. Remember, in our previous discussion on dissolving the problem I mentioned Skorupski's bridge-principle that went as follows:

> *The Bridge-Principle*: If an agent, A, has a reason to have attitude *x* then, and because of that, A has a reason to do *x*-prompted actions (Skorupski, 2010b, p. 129)

The bridge-principle is a first attempt to find a connection between the evaluative and reasons. Furthermore, it is intuitively the case that most of our evaluative reasons to blame an agent stem from practical reasons the agent had to act otherwise. Most of evaluative the reasons to blame that we normally cite refer to practical reasons the agent had – e.g. "The agent had promised to do so but did not keep her promise".

From this discussion of supererogation, I think we can see that the notion of supererogation is the basis of a good argument against O-OB, and that it is more plausible to reject O-OB than OB-Blame. Moving on, a cost of accepting FA-Blame is that by accepting an FA analysis of obligation we will have to deal with all the problems raised by the FA analysis, including the difficulty with reasons of

---

[95] For more on consequentialism and supererogation, see (McConnell, 1980; Portmore, 2003; Vessel, 2010).

the wrong kind to blame, and so on. Suppose, for instance, that an evil demon threatens me with torture unless I blame you if you do not clap your hands five times when I say so. Does that mean that I have a (conclusive) reason to blame you if you fail to obey me, and that you therefore are obligated to do so, even though it is not the case that you ought to do it?

Finally, I want to say a few words about *MOB-Blame* and OB-Blame. *MOB-Blame*, while formally silent on the matter, surely carries the implicature that there can be non-moral obligations that we are not blameworthy for violating. If this is correct, it seems that FA-Blame is at most an analysis of *moral* obligations, and that we have not given an account of *obligations* in general. Furthermore, if blame is an inherently moral notion, then it seems that accepting FA-Blame as an account of obligations in general implies that all of our obligations are moral obligations. To my ears, this does not sound strange at all. It is, however, quite committal.[96] To be less committal, we might want to generalize FA-Blame to include other holding-accountable attitudes.

> *FA-OB*: For it to be the case that *A* is obligated to *φ* is for it to be the case that there is (conclusive) reason to have a holding-accountable attitude towards *A* if *A* does not *φ*.

If it turns out that all our obligations to do something are moral obligations, then perhaps the set of 'account-holding attitudes' will only include blame and other moral sentiments. If, on the other hand, there are non-moral obligations, the set of account-holding attitudes might include other attitudes. For completeness' sake, let me spell out how the FA analysis accounts for the other deontic properties of being permissible, being impermissible, and being optional.

> *FA-PE*: For it to be the case that *A* is permitted to *φ* is for it to be the case that there is not (conclusive) reason to hold a holding-accountable attitude towards *A* if *A* does not *φ*.

---

[96] How committal this actually is will depend on what one means by 'moral obligations,' and on how all-encompassing the sphere of morality is. If morality has a very wide scope of application, it may not be so committal to claim that all obligations are moral. If the scope of morality is narrow, it will sound counterintuitive to say that the only obligations we actually have are moral ones. I thank David Alm for forcing me to think about this.

FA-IM: For it to be the case that *A* is forbidden to *φ* is for it to be the case that there is (conclusive) reason to hold a holding-accountable attitude towards *A* if *A* *φ*s.

FA-OP: For it to be the case that *A*'s *φ*ing is optional is for it to be the case that there is neither (conclusive) reason to hold a holding- accountable attitude towards *A* if *A* does not *φ* nor (conclusive) reason to hold a holding-accountable towards A if *A φ*s.

To summarize, I believe that there is good reason to accept the Standard Account, passing the ought-buck to practical reasons. I think supererogation gives us good reason to reject O-OB, and thus that what we are obligated to do is not the same deontic property as what we ought to do. Instead we should pass the buck from OB, PE, IM, and OP to evaluative reasons to blame and other holding-accountable attitudes.

Ought Implies Can and Reasons Imply Can

One of the better-established principles in moral philosophy is that if an agent ought to *φ*, it follows that the agent can *φ* – the familiar slogan is 'ought implies can' (OIC). OIC can be used to mount an objection to the Standard Account.[97]

OIC is meant to explain why one cannot be morally required to jump to the moon, change the past, physically save a child in Paris, this instant, when one is in London, or perform two incompatible actions (as in Sophie's Choice).[98]

---

[97] OIC is usually attributed to Kant and is sometimes referred to as Kant's maxim (Copp, 1997, 2003). For interpretative issues dealing with how Kant himself viewed OIC, see (Baumgardt, 1946; Stern, 2015).

[98] Sophie's choice concerns what is usually referred to as a "tragic dilemma". In the choice situation, Sophie can save one of her two children, and if she does not make a choice they will both die. The discussion in the literature is whether there are any genuine dilemmas, i.e. situations where we are morally obliged to do two or more incompatible actions. One line of reasoning is that Sophie ought to save her first child, child A, and she ought to save her second child, child B. On the principle of *agglomeration* in deontic logic, if one ought to ought to *φ* and one ought to *φ**, it follows that one ought to *φ and φ**. This leads, however, to violation of OIC. For this reason, some reject OIC, and others deny agglomeration (either fully (Bergström, 1977) or restrictedly (Horty, 2012)). In the Sophie's choice scenario, it has been denied both

If we can understand what we ought to do in terms of what we have conclusive reason to do, it cannot be that we have a reason to act that would violate OIC. In other words, if we want to pass the deontic buck and accept OIC, we are committed to certain ideas about how to formulate the principle 'reasons imply can' (RIC). I shall argue that this need not worry us – that we should be willing to interpret RIC in the same terms as OIC, but that OIC (and RIC) might need adjusting when we are dealing with evaluative, or theoretical oughts (and reasons).

Peter M. Vranas, in his influential paper "I Ought, Therefore I Can" (2007), argues for OIC on the grounds that OIC is entailed by RIC and RIC is obviously true. Vranas presupposes that RIC is valid, and that we should understand RIC and OIC in the same way. I will explore some of the other possibilities here.

First, as long as one accepts the Minimal Thesis (see above, page 79) that whenever an agent ought to do something then the agent has a *pro tanto* reason to do it, it will follow that RIC, at least, cannot be more restrictive than OIC: It will follow, that is to say, that whenever an agent 'can' according to OIC, that agent 'can' according to RIC. Given that the Minimal Thesis is almost universally accepted (again, *pace* Horty (2012) I shall assume that the only interesting question is whether the 'can' in RIC is less restrictive than the 'can' in OIC.

Once we move beyond OIC's slogan nothing is uncontroversial. First, it is controversial whether OIC is a substantial or a meta-normative principle. Can OIC be denied within a first-order theory about normativity? Or does it say something conceptual about 'ought'? Secondly, it is a matter of debate whether OIC applies only to moral oughts or to other types of ought as well (Wedgwood, 2013; Fox & Feis, 2018). Thirdly, it is contested whether 'implies' should be understood as an entailment relation. Should we say instead that the 'can' is merely conversationally implied (Sinnott-Armstrong, 1984; Hieronymi, 2004; Vranas, 2007)? Fourthly, in the debate it has not yet been made clear what is meant by 'can'. Does it refer to physical possibility, or should we understand it, along lines suggested by Frances Howard-Snyder, as meaning "that it is highly likely that the agent successfully does it" (Howard-Snyder, 1997)? Or should we follow Peter Vranas' approach, and take 'can' to mean "that the agent has the

---

that Sophie ought to save A and that Sophie ought to save B, and insisted that what she ought to do is save A *or* B.

opportunity to do it and the general ability to do it, even if success is not guaranteed or even likely"(Vranas, 2007).

In Paper III, "A Fair Reading of 'Ought Implies Can'", I contend that in moral cases we should take the 'can' in OIC to mean 'if the agent tried her best to do it, she would do it'.[99] I argue that any other understanding leads to unfair moral requirements, and that moral requirements cannot be unfair (Copp, 2003). Intuitively, it would be unfair, in cases where the agent has tried her best to do something, to claim that she failed to do what she was morally required to do.

The strongest alternatives to my own proposal as to how we should interpret the 'can' in OIC include the following (further discussion of these is presented in Paper 3):

> *Streumer's proposal*: An agent, *A*, can φ if and only if there is a nomologically and historically accessible possible world where *A* φs (Streumer, 2007).

> *Vranas' proposal*: An agent, *A*, can φ if and only if *A* has the general ability to φ and the opportunity to φ (Vranas, 2007).

> *Howard-Snyder's proposal*: An agent, *A*, can φ if and only if *A* is very likely to φ if A tried to φ (Howard-Snyder, 1997).

The proposals are ranked in terms of how permissive they are. Bart Streumer adopts the most lax understanding of 'can'. My proposal is the most restrictive of the options. For present purposes, however, we need not commit to a specific reading of OIC.

Streumer (2007, 2010, 2018) argues for the thesis that *is* such a principle as RIC.[100] I find his arguments persuasive, but he only argues for the claim that there

---

[99] There are, of course, other senses of 'can' – there are senses of 'can' in which I can jump to the moon, beat Magnus Carlsen at chess, and make a 100% of my basketball shots. The sense of 'can' that we are now discussing is the sense relevant to moral requirements.

[100] For someone who has argued for the existence of reasons to do what we cannot do, see Ulrike Heuer (2010). Heuer believes that we sometimes have derivative reasons to do what we cannot do, such as learn to play the piano. She further argues that the only way we can get these derivative reasons is by having non-derivative reasons to do the thing we cannot do. By contrast, Kimberley Brownlee (2010) utilizes the notion of *ideals*. An ideal (e.g. being the perfect

is a principle of RIC. This leaves it open to question whether there is still a difference, however slight, between OIC and RIC. Of course, the Standard Account could be seen as evidence that there cannot be even a slight difference between OIC and RIC. Against this, it might be argued, however, that it seems that evaluative reasons and theoretical reasons might not be constrained in the same way, by ability, as what we ought to do. For instance, it is possible I have reasons to believe that p even if it is impossible for me to actually believe that p because I have taken a drug that makes me always believe not-p – or because I do not have sufficient control over my own belief formation. The same thing might be true of reasons for attitudes. We can imagine a person who is incapable of having certain attitudes but still has reasons to admire Derek Parfit.

One could argue that RIC should be the same principle, *mutatis mutandis*, when applied to practical reasons, theoretical reasons and evaluative reasons. But note, however, that it might just as well be true that OIC takes different form when applied to what we ought to believe, or feel, and what we ought to do. This last alternative looks quite plausible, so the objection from OIC should not worry deontic buck-passers.

Taking Stock

In this chapter I have discussed how reasons relate to other normative phenomena. I have focused especially on the passing of two kinds of buck: The *evaluative* buck and the *deontic* buck. It cannot be denied that there is more to say on these issues, and that the buck-passing accounts presented have their fair share of objections to deal with. I believe, however, that the hurdles they face are manageable, and that they pale in comparison to the theoretical gains of the accounts. Added to that, they manage to be very informative while remaining theoretically parsimonious. All in all, these are exciting times for normativity research. In the next chapter I will ask how different sources of reasons and domains of normativity relate. In short, I will offer a reason-based account of the architecture of normativity.

---

violinist) is by definition impossible to reach. Nonetheless, according to Brownlee, we can have reasons to match an ideal (to become the perfect violinist).

How It All Relates

This thesis is about the structure of reasons and normativity. As I said on page 16, I approached this extended introduction with two aims. The first was to set out the background to the papers, and to situate them within a wider philosophical context. The second was to provide an overview of some of the concepts and properties which play a prominent role in the literature on the nature of reasons and normativity. I believe that these two aims have been largely achieved.

We have discussed what it is to be a normative reason. We have distinguished between *pro tanto* reasons and overall reasons. We have critiqued the weighing metaphor and the way in which reasons are meant to be weighed. We have looked at how reasons relate to values and certain deontic properties. With these discussions behind us, the papers will be a breeze.

One topic we have only briefly touched on is that of different kinds of reason, such as moral reasons and prudential reasons. While coverage of this topic may not be necessary for the papers per se, I none the less want to conclude this extended introduction with a few structural remarks about kinds of reasons, and how this topic relates to the wider architecture, or structure, of normativity. I believe this will neatly tie the previous discussions together.

Some parts of this chapter proceed at a high level of abstraction. For this reason, I will be utilizing toy theories to illustrate some of the different points I will be making. These toy theories might not be plausible normative theories, but that is not their aim. Their aim is to be clear and simple. While the chapter operates at a high level of abstraction, the basic questions we are trying to answer are relatively simple: There seem to be different types of reasons for different things – but what different things? There seem to be different kinds of reasons – but what are these kinds? What are we trying to convey when we say that something is not just a normative reason but, say, a moral reason? Can our reasons to believe clash with

our reasons to act, and if so, is there anything meaningful we can say about how to resolve the conflict?

What will emerge from the answers to these questions is what I, following (Reisner, 2015b), call theories of the architecture and structure of normativity.

In the first chapter we touched on an answer to the first question. I claimed that there are three types of reasons: Practical, theoretical, and evaluative. We can see how each kind corresponds to a different interpretation of the φ-place holder in the basic reason relation R(x, S, φ). Each interpretation designates a normative domain.[101] With our three different interpretations of φ we arrive at three domains of normativity: The domain of practical normativity, where φ is an action, the domain of theoretical normativity, where φ is interpreted as a belief, and the evaluative domain, where φ is an attitude (i.e. a feeling, or emotion, or other non-doxastic attitudes). This thesis can therefore be said to present a pluralist view of normative domains. I will now present a few ideas about domains of normativity, and how these domains can relate. I will then move on to discuss some of what we commit ourselves to by saying that reasons are of different kinds, and finally I shall examine different ways to deal with conflicts between reasons from different domains and between reasons of different kinds.

Normative Domains

As has just been said, I assume a plurality of normative domains. Of course, pluralism about the domains of normativity is not necessarily confined to the three domains I have indicated; one can imagine quite a few varieties of pluralism about normative domains. A common position is the dualist view that there is a theoretical and a practical domain of normativity. We can imagine other candidate domains: The domain of intentions, or perhaps that the domain of emotions and desires, may be regarded as separate domains, and so on. However, as long as one accepts more than one normative domain, one is a pluralist.

The next question to ask oneself in shaping the theory of the architecture of normativity is whether reasons (and oughts) from different domains can come

---

[101] For a defense of 'triplism' about normative domains, see (Skorupski, 2010a, 2012).

into conflict. According to Reisner, putative inter-domain conflict occurs "when what you ought to do according to one domain precludes you from doing what you ought to do in another domain" (Reisner, 2015b, p. 192). He goes on to give the following example:

> Suppose that the evidence strongly suggests that *p* is the case. On a number of common views about theoretical reasons, one ought to believe *p*. Suppose, however, that believing *p* will make you depressed and will interfere with your life. […] On a number of common views about practical reasons, you ought to cause yourself to not believe *p* (Reisner, 2015b, p. 192).

Since you cannot successfully both believe *p* and cause yourself to not believe *p* there is, according to Reisner, an inter-domain conflict. If two oughts from distinct domains are in conflict with one another, it seems to make sense to ask:

> *Conundrum*: So, what ought I to do in the end? What I ought to do relative to the theoretical domain, or what I ought to do relative to the practical domain?[102]

What one seems to be asking is: Which ought is more important? The practical one or the theoretical one? To answer this question, we need a way to compare the relative importance of the two oughts. This, in turn, requires us to identify a respect to which we can measure the importance.[103] In other words, there needs to be an overarching domain, an overarching ought that subsumes the theoretical and practical domains, if the comparison is to make sense.

To believe that there is a meaningful answer to Conundrum is to adopt the position Reisner (2015b, p. 193) labels *normative non-separatism*. A normative non-separatist needs a story about how to settle inter-domain conflicts, or about why putative conflicts are merely apparent rather than real. As regards the settling of inter-domain conflicts, a normative non-separatist can claim either that there

---

[102] Reisner formulates a similar question in terms of "Ought I to do as practical rationality says that I ought, or as theoretically rationality says that I ought?"(2015b, pp. 192-193). In the quoted question he has a footnote which states that the "do" is supposed to be read as a universal verb "meant to include beliefs and feelings" (Ibid).

[103] Following Andersson (2017), what I have in mind here is that all comparisons must be made relative to a 'covering concept' – e.g. coffee is more important than toast relative to Jakob having a nice breakfast. In the case of judging the importance of the two oughts there needs to be a covering concept according to which we can compare their degree of importance.

is a way to determinably weigh reasons from different domains, or that the weights of reasons from different domains are incommensurable.

The alternative is to accept *normative separatism* and deny that the Conundrum question has a meaningful answer. Relative to the theoretical domain you ought to 'believe *p*', and relative to the practical domain you ought to 'cause yourself to not believe *p*' – end of story. The theoretical reasons for believing *p* cannot be weighed against the practical reasons to cause yourself to not believe *p*. According to a normative separatist, there are domain-specific reasons; there is one ought for each domain, but the domains are wholly distinct. There is no further question to ask – reasons from different domains are not only incommensurable, they are *non-comparable*. According to (Andersson, 2017), two objects are non-comparable if there is no covering concept that is applicable to both objects. For example, the number 7 and blueberry pie are non-comparable in terms of tastiness. Blueberry pie is not tastier than the abstract number 7, because that abstraction is not something to which the covering concept of tastiness can be applied.[104] A normative separatist can either deny that there are inter-domain conflicts or accept that the relevant inter-domain conflicts are unsolvable.

I believe that there is something suspicious going on in Conundrum. The suspicious aspect is imported by the use of 'do' as a universal verb. Of course, I have been using 'do' as a universal verb throughout most of this thesis (see page 15 footnote 3). My intention, however, has been to use it as a placeholder for an action, belief, or attitude. According to Reisner (2015b, p. 192fn) no universal verb exists in English. Hence, he invents it and stipulates that he will use 'do' as a universal verb. My suspicion is that the universal verb 'do' may be confused, unless it is merely used as a placeholder. Consider:

---

[104] It might be useful to compare this with a case of conflict within a domain. Assume that your house is on fire and you can save either the manuscript of your superb novel or your boyfriend. Aesthetically speaking, it may be that you ought to save your manuscript. Morally speaking, it may be that you ought to save your boyfriend. In this scenario, it might make sense to ask: In the end, which action ought I to perform? There is a covering concept, so to speak, where ideally you can compare the importance of the two oughts. Relative to the practical domain, what is more important in this scenario – morality or aesthetics? In the case of inter-domain conflict, I wonder whether there really is a sensible conception of an overarching domain such that a comparison between a practical ought and a theoretical ought makes sense.

Conundrum*: So, what ought I to believe/feel/do in the end? What I ought to believe or what I ought to do?

Here "do" is not used as a universal verb. It explicitly designates the performance of some action. In Conundrum*, if "believe/feel/do" is interpreted as a placeholder, so that we have to fix the interpretation of the placeholder either to "believe", or to "feel" or to "do" before answering the question, the answer to the question (in each case) becomes obvious. If we interpret the placeholder as an action, then you ought to perform the action you ought to perform. If we interpret the placeholder as a belief, then you ought to believe what you ought to believe. It is only if we interpret "do", not as a placeholder, but rather as an actual universal verb, or as the conjunction "What ought I to believe-feel-do?", that we run into trouble.

This is a mere sketch of an argument against the idea of a proper universal verb, i.e. to believe-feel-do makes sense. But if it is sound, it might work against the idea of an overarching domain. If it were sound, then, assuming there are inter-domain conflicts, they would, by necessity, be unsolvable conflicts. I am unsure if this reasoning is cogent, but if it is, then, it is an argument in favor of normative separatism.

In addition to needing to posit an additional overarching domain, I think there might be another reason to hesitate before accepting normative non-separatism. I believe that there might be an important difference between normal (genuine) dilemma situations, such as *Sophie's choice*, and putative inter-domain dilemmas. In the standard set up, for a *Sophie's choice* dilemma to count as a genuine dilemma it must be the case that the you ought to rescue child *A* and ought to rescue child *B* but cannot rescue both *A* and *B*. In the *Sophie's choice* scenario, another valid description of the action described by the phrase "rescuing child *A*" is that you are "not rescuing child *B*". Because of this, we can describe the *Sophie's choice* scenario as a case where you ought to φ and ought to not-φ. Rescuing child *A* entails not rescuing child *B*, and vice versa. It is a genuine dilemma, not only because it asks us to do something that we cannot successfully do, but also because it asks us to perform a contradiction.

In the case of inter-domain conflict – say, where you ought to believe *P* and you ought to cause yourself to not believe *P* – I am unsure if the entailments are equivalent.

It is perfectly true that believing *P* implies that you do not (successfully) cause yourself to not believe *P*. I am less sure, however, that not causing yourself to believe not *P* implies that you believe *P*. It might be true that you do not cause yourself to not believe *P* while, at the same time, it is also true that you believe not *P* or lack a belief in either *P* or not *P*. This might give us reason to hesitate to describe the inter-domain conflict as a case where you ought to *φ* and not-*φ*.[105] This might lead us to question whether there is a genuine conflict between the two domains. Is it, rather, merely a tragic situation?[106]

Normative Sources

We often theorize about different kinds of reasons. We talk about what moral, prudential and aesthetic reasons we have. Some might say that my promise to my friend to read one of her papers is a moral reason to do so. They might add that the fact that I could work on my own paper instead of reading hers is a prudential reason (not a moral reason) to not read her paper. Finally, some might say that my friend's eloquent style of writing is an aesthetic reason to read her paper.

The modifications 'moral', 'prudential' and 'aesthetic' signal the type of reason in play. What type of reason is in play depends on what the normative source of the reason is.[107] More generally, we can say that in the following propositions "*x* is an

---

[105] Furthermore, it is quite plausible that belief formation is involuntary, or at least that belief formation is under less control than action. For this reason, OIC might differ substantially between the theoretical domain and the practical domain. It seems possible that, from the point of view of the theoretical domain, the tragic situation that you cannot successfully believe *P* and cause yourself to not believe *P* does not violate the OIC of the theoretical domain since we are not dealing with a logical contradiction. My thanks are owed to Marta Johansson for fruitful discussions of this issue.

[106] Note that (as we will be discussing below) it is possible to contend that there are both evidential and pragmatic reasons for beliefs. If that were the case, there would plausibly be intra-domain conflict within the theoretical domain over whether to believe *P* or not-*P*, but such conflict would be structurally very different from inter-domain conflict over whether to believe *P* or cause yourself to not believe *P*.

[107] The distinction between kinds of reasons (practical, theoretic and evaluative) and types of reasons (moral, prudential, aesthetic, and so on) is orthogonal in that it is at least a logical possibility that there are, for example, moral evaluative reasons and aesthetic practical reasons.

N-reason for *S* to *φ*" and "*S N*-ought to *φ*" the '*N*' signals the type of reason or ought. In other words, *N*, in this framework, signals the *source* of the reason (or ought).

In my view the concept of a 'source of reasons' is intuitive in the sense that most competent English speakers can use it in natural conversation. By 'source of reasons' I mean, roughly, that partly in virtue of which a certain fact has the normative property of being a reason.[108] To illustrate, take the fact that Ann promised Beth to *φ*. It is something about the nature, or essence, of promises that partly explains (i.e. grounds) the fact that Ann's promising Beth to *φ* has the property of being a reason for Ann to *φ* – it has fidelity as its source.[109] Fidelity can plausibly be taken to be a sub-class of both morality and prudence.

However, like Broome (2013) and Reisner (2015b), I cannot, in Broome's words, "say exactly what sort of thing sources of requirements [or reasons] are – what category the law, morality, rationality, prudence, convention, fashion, etiquette, honour, the rules of chess and all the rest fall under. I do not know a generic term for things of this sort" (Broome, 2013, p. 116). As can be seen from the list Broome presents, not all sources need be truly normative in the sense spelled out on (page 19 and on page 81). For instance, I might have an etiquette-based reason not to eat the main course with the desert fork, but according to most normative theories, this reason is (usually) not a truly normative one. Our focus is solely on proper normative sources of reasons.

Furthermore, if one accepts any of morality, prudence or aesthetics as normative sources, it is quite likely that one will view these as umbrella notions – notions covering sets, or collections, of sources. I refer to sets of this kind as 'normative spheres'. Make no mistake, normative spheres are normative sources, but I believe

---

[108] I use the phrase "partly in virtue of" to signal that it is not only in virtue of the normative source that a certain fact has the normative property of being a reason. The fact in question has the property of being a reason partly in virtue of its natural properties *and* partly in virtue of certain background conditions, as was discussed in on pp. 48-50), and as is further elaborated on in Paper I and Paper II.

[109] My usage of 'source' is inspired by Broome's (2013, p. 116), with small modifications. For Broome, a source of reasons is a requirement, an ought, that explains why a specific fact is a reason. For me, it is also true that the source explains (or grounds) why a certain fact has the property of being a reason, but the source is not a requirement. For more on sources, reasons and requirements, see (Guindon, 2016).

it is good to have a vocabulary to distinguish complex sources such as 'morality' from simpler sources such as 'wellbeing'. A normative sphere is simply, by my stipulation, a collection of normative sources.[110] To illustrate, I want to introduce a toy theory about the sphere of morality:

> *Toy theory 1*: There are two domains, the theoretical and the practical. There is one sphere, morality, which in turn consists of two sources, fidelity and wellbeing.

According to our toy theory, all moral reasons are either promise-based or welfare-based. On the theory, morality is a normative sphere. Toy theory 1 says nothing about how to settle conflicts between reasons arising from fidelity and reasons arising from wellbeing. We will talk more about intra-sphere conflict shortly.

One can be either a pluralist or a monist about sources. Source monism is the view that there is only one source.[111] An example of this kind of view, borrowing from Reisner (2015b), is a radical pragmatist, i.e. someone who views wellbeing as the only source of reasons, be those reasons for belief, action or what have you. As we can see from the radical pragmatist, a specific source can function as a source for a single normative domain or several normative domains.

---

[110] For the thoughts of someone who has written extensively on spheres of normativity – or rather, "points of views" as she calls them – see Susan Wolf (2015), especially Chapter 4.

[111] Reisner (2015b, p. 197) uses the labels "normative monism" and "normative pluralism". I prefer "source monism" and "source pluralism" because I believe that "normative pluralism" and "normative monism" are too ambiguous without a proper context. For instance, a reader might take those labels to refer to monism or pluralism about fundamental normative properties, such as reasons-first or values-first views. I recognize, on the other hand, that my preference invites the objection that the labels are not informative enough without a proper context.

Normative Sources and Conflict

Just as with normative domains, a decent heuristic for finding out the number of sources a normative theory has is to look at the reason relation, R(x, S, φ).[112] In order to determine the number of sources, we look at the x-.variable. Depending on the different varieties of fact the x can be ('varieties' in the sense given by the question: 'What is the fact about?') we get a rough idea of how many sources a theory posits. Now, this is a bit tricky, because very often the source, or the type of fact, is not immediately recognizable, or readily distinguished. At times this will be easy. For example, if the reason, x, is the fact that "Ann promised Beth" it is easy to see that it is very likely to be a reason with fidelity as its source. At other times, however, it will be harder to detect the source, as is shown by an example given by Reisner:

> The fact that eating sweets will make you happy is a reason for you to eat sweets. The normativity of that reason is basic. Eating sweets is one of the things that make you happy. Saving money, on the other hand, does not make you happy. However, in order to buy sweets, you must have saved some money. So there is a derivative reason to save your money. The fact that saving money will allow you to buy sweets, which in turn will make you happy, is a reason for you to save your money (Reisner, 2015b, p. 197).

In this case, welfare (or happiness) is the normative source that explains, in part, why the fact that saving money will allow you to buy sweets has the property of being a reason for you to save money. This is so, even though the reason-fact is not explicitly 'about' welfare. It is a derivative reason, i.e. a reason that is a reason in virtue of something else that is a reason.[113]

---

[112] An area of interest bypassed in this thesis is the issue of agent-relative and agent-neutral reasons. This issue, which has been widely debated, is not of immediate concern for the papers in the thesis. For this reason I have chosen to stay away from that particular hornet's nest. I remain skeptical as to whether there is a useful distinction between agent-neutral and agent-relative reasons. If there is a distinction to be made, I believe that a fruitful way to characterize different ways to draw the distinctions as different understandings of permissible interpretations of the agent-variable, S, in the reason-relation. For more on agent-relative and agent-neutral reasons, see (Nagel, 1970; Rønnow-Rasmussen, 2009; Buckland, 2017).

[113] For more on derivative reasons, see Paper IV. There I defend the view that if you have a reason with a certain weight to perform some action φ, and ψing is a necessary means of φing, then,

Once we have adopted a view on the plurality of normative sources, we need to develop an account of the way these sources relate to domains. Is there a conceptual connection between particular sources and particular domains? For example, is there a conceptual connection between evidence and the theoretical domain that is lacking in the other domains, so that evidence is only ever a reason for belief and never a reason for action?

Structurally, there are two types of position (Reisner, 2015b, p. 198). One either holds the view that (at least) some sources are conceptually linked in a satisfactory manner to some specific domains, or maintains that there are no such conceptual connections between particular sources and specific domains. On the latter view, any source could, potentially, provide reasons for all domains. Reisner labels these two different types of view 'reasons specialism' and 'reasons generalism'. Reasons specialism is the view that there is a conceptual connection between certain sources and certain domains; reasons generalism is the denial of reasons specialism (Reisner, 2015b, p. 198).[114] Exactly what it would take for a conceptual connection to be 'satisfactory', however, is regrettably something that is outside the scope of this thesis and I have to leave it as a rather vague notion.[115]

---

and because of that, you have a reason of the same weight to ψ. The same applies to what you ought to do – if you ought to φ, and ψing is a necessary means of φing, then, and because of that, you ought to ψ.

[114] It could turn out that a reasons specialist and a reasons generalist are *a posteriori* in agreement on which sources *in fact* provide reasons for which domains but still disagree *a priori* on why this is the case (Reisner, 2015b, p. 200). The reasons specialist and reasons generalist may both accept that *in fact* only evidence functions as the source of reasons for belief. They could still disagree over whether there is a conceptual connection that evidence shares with reasons for belief and that other sources of reasons do not.

[115] I am not in the business of spelling out the actual conceptual connection between different kinds of reasons and different types of reasons, i.e. explaining what the conceptual connection between sources and domains might be. I should have hoped that I have at least dabbled in the question of what type of demands we could expect from a satisfactory conceptual connection. Sadly, I do not know what demands we should expect from a satisfactory conceptual connection. That may depend on a wide range of things, depending on the domain and the source in question. For discussion of a potential conceptual connection between evidence and the theoretical domain that would also rule out the possibility of any other sources being reasons for the theoretical domain, consult the evidentialist (Shah, 2006). For objections to Shah's view, see (Yamada, 2010; Sharadin, 2016).

To illustrate what has been said so far, let us elaborate Toy theory 1. Toy theory 1 is the claim that the sphere of morality consist of two normative sources, fidelity and wellbeing. Let us now expand this theory:

> Toy Theory 2: There are three domains, the practical, the theoretical and the evaluative. The practical domain is conceptually connected to one sphere, morality. Morality consists of two sources, wellbeing and fidelity. The practical domain has no other normative sources. The theoretical domain is conceptually connected to one source, evidence. The theoretical domain has no other sources. The evaluative domain is conceptually connected to the sphere of morality, and the sphere of aesthetics. The sphere of aesthetics consists of two sources, sublimity and originality.

This theory has three domains, two spheres, and five sources. Since some sources are sources of reasons for some domains but not all, Toy theory 2 is an instance of reasons specialism. It is pluralistic about sources, and pluralistic about normative domains.

When we are characterizing the architecture of normativity, a vital question – and perhaps the question most thoroughly discussed in the literature – is how to handle putative intra-domain conflict. The source monist has a relatively easy story to tell about intra-domain conflict. Of course, a source pluralist might claim that for some domains there is only one source. For instance, utilitarians who believe that the only source in the practical domain is utility have an easier time dealing with conflict within the practical domain. They just need to weigh up how much utility is associated with each alternative.

On the other hand, the fact that all the reasons derive from the same source does not show that the process of weighing up these reasons will be simple. Depending on what the sources are, there could be issues raised by organic unities, or problems with incomparability – although, as discussed in (Chang, 2002a; Broome, 2004b; Rabinowicz, 2008, 2012; Temkin, 2012; Andersson, 2016, 2017; Elson, 2017), the culprit of incomparability is often in one way or another that there are several aspects that needs to be weighed together. With only one

source it is likely that the problem of having to way together several aspects will be less of an issue.[116]

For someone who accepts that there is more than one normative source in a domain, there are various ways to handle intra-domain conflict. Conflict between spheres and sources in the practical domain is an age-old issue. For example, does morality always trump prudence and aesthetics? Some theorists may wish to argue that the spheres display a certain intra-domain hierarchy, whereby, for example, moral reasons strictly trump all other spheres in the practical domain. Similarly, it is a contentious topic how robustly wellbeing-based reasons stand up to fidelity-based reasons. If we follow Kant, we will get one answer, but following (Ross, 1963/1939, 2002/1930), we may conclude that no systematic answer can be given to the question of how to solve these conflicts.

It could be that there are different ways to solve intra-domain conflict, depending on the domain. It is no small task to explain how intra-domain conflicts, or even intra-sphere conflicts, are supposed to be handled. Explicating the sources of morality, and how different sources are supposed to compare, and whether they can come into conflict, is equivalent to the entire field of first-order normative theory. At the start of this thesis (see page 21), I made it a point of pride that I would be steering clear of first-order issues and conducting discussion at the structural level. For this reason, I am not arguing for any particular theory when it comes to any of the issues discussed.[117]

---

[116] This is not to say that they will not occur. Take for instance wellbeing, we have different qualities of wellbeing, different intensity and temporal duration and so forth. For more on hard comparisons and incomparability see (Andersson, 2017, pp. 108-115)

[117] There are, however, a few things that can be said at the structural level about substantive normative theory. I believe that any substantive theory describing a domain, or sphere, can be usefully represented by two parameters: (i) a specification of the normative sources, and (ii) a specification of a weighing function and the way the sources relate. In other words, a substantial normative theory needs to specify what matters, and how it matters. In any given situation, a moral theory assigns certain facts the property of being a reason for an agent to do something, and tells us how to weigh these reasons against each other. Recently, Franz Dietrich and Christian List (2017) have offered a framework on which a very wide range of substantial moral theories can be represented in terms of these two parameters. Although Dietrich and List focus on moral theories, I see no reason why their general methodology and approach could not be generalized to any normative sphere, or any normative domain. For more on their elegant formalization and some structural results, see (Dietrich & List, 2017).

What emerge from this discussion are several possible normative architectures with rather different structural features. Combining these features in different ways creates contrasting theories of the architecture and structure of reasons and normativity.

With these theories, we can note that we face a methodological dilemma. Different architectures have different consequences for the first-order theories. We can see how "substantive claims about oughts and reasons limit (or are limited by) the general structure of a normative theory" (Reisner, 2015b, p. 204). The methodological dilemma is whether to work from the bottom up, the top down, or in some sort of dynamic way. Should theories of the architecture of reasons and normativity be seen merely as descriptions of what the different normative claims amount to, or do they have a prescriptive element? If some theories are better than others, maybe we should disregard first-order normative theories that imply such architecture.

Conclusion

In this extended introduction, I have presented an exposition of some of the current debates revolving reasons and normativity. I have done so with the focus on addressing structural issues and providing some sort of a cohesive story about some of the more salient concepts employed in the papers to follow. Along the way, I have offered an overview of some of the contemporary themes within the research field, such as how to understand the property of being a reason, the fitting attitudes analysis of value, the connection between what we ought to do and what have reasons to do, and how our obligations depend on reasons.

In the thesis we can see several points where there is a need for more research. I shall indicate a few such points. A lot more needs to be said about the weighing metaphor. More research is also needed into whether the FA analysis can be fruitfully extended to morality. All of the potentially different ways in which the attitudinal component and the normative component of the FA analysis affect the value of an object need to be teased out. A particularly interesting topic is whether we can pass the buck to other philosophically interesting concepts. An interesting concept that I have in mind here is probability. This has already been attempted (Skorupski, 2010a; Rabinowicz, 2017), but I think that an extension of Paper V

will show that more needs to be said on the endeavor and its advantages. On the topic of the FA analysis, research is needed on whether there is a philosophically interesting distinction to be drawn between the notion that it is fitting to have a certain attitude and the notion that it is fitting to express the very same attitude. For instance, it is my impression from reading the literature on 'sufficient standing to blame', that we are not differentiating, to a satisfactory degree, between an attitude being fitting to have and an attitude being fitting to express. The fact that an attitude is fitting in the first sense does not guarantee that it is so in the second. These are just a few of the areas which this thesis suggests stand in need of more research. The philosophical interest in reasons and normativity has started to pick up speed only in the last twenty-odd years and I believe that this field of research is still in its infancy.

The papers constituting this thesis present in-depth discussions of certain aspects and issues concerning reasons and normativity. I shall now provide a brief overview of each paper.

Overview of Papers

The five papers that follow are self-contained and should be read as such. The order of appearance is thematic in terms of generality, from very general questions about normativity to specific questions about how the Fitting Attitudes analysis of value should solve a specific problem. The co-authored papers are a product of mutual and equal labor. The advantage of sharing office with my co-authors has resulted in the papers having been written in tandem, sitting side by side. The order of authors is alphabetical. The papers under review are presented here as they were submitted with the exception that redactions for reasons of anonymity have been restored. Paper I presents an account of how to understand normative background conditions in a non-deflationary manner. Paper II deals with how the distinction between subjectivism and objectivism about value and reasons is best understood and how to give a positive interpretation of the notion of constitutive grounds. Paper III deals with how we are to best understand 'can' in the dictum 'ought implies can'. Paper IV defends the idea that reasons, and oughts, transmit fully from actions to the necessary means of that action. Paper V deals with how the Fitting Attitudes analysis of value should understand degrees of value.

In what follows, I provide a brief summary of each paper.

Paper I. Towards A Reasons-First View of Background Conditions

In this paper Andrés Garcia and I develop an account of background conditions for normative features, such as enablers and disablers. In chapter 1 I briefly discussed the idea that normative features are subject to background conditions.

The central thesis of the paper is that if we assume a metaphysical hierarchy of normative features then there is nothing really surprising or mysterious about background conditions. We argue that our account is better suited to account for the intuitions surrounding cases of background conditions than its pragmatic

alternative. The developed account also has the advantage of being general in the sense that it can explain background conditions for all normative phenomena.

The first part of the paper makes the case via examples that there is an interesting phenomenon that calls for inquiry. By looking at the previous literature on the issue we establish two functional criteria to demarcate what a theory of (normative) background conditions need to conform to. Roughly, in order for a factor to be a background condition for *value* it needs to be a part of what explains the instantiation of the value but not be a value-bearer or a value-maker. For a feature to be a background condition for *reasons* it needs to be a part of what explains why that reason is instantiated without the factor itself is not a part of the reason itself.

The second part of the paper claims that if we assume a Reasons-First approach to normativity, or at least assume the FA analysis of value (see chapter "On Buck-Passing"), we can identify features which align with the functional criteria laid out in the first part. On a Reasons-First approach, reasons function as value-makers. In the paper we argue that the features in virtue of which certain facts are endowed with the normative property of being a reason, i.e. the reason-makers, can fruitfully be identified as the background conditions for value. This can be done in a way that both explains our intuitions about specific cases and conforms to the functional criteria. It does so without introducing any hitherto unknown ontologically extravagant *sui generis* relation. The functional criterion for background conditions for reasons was that background conditions are not part of the reason itself, and hence the background conditions for reasons are identified with those reason-makers that are extrinsic to the reason-fact.

The third part of the paper investigates the advantages our approach has in comparison to other alternative explanations, such as the pragmatic deflationary understanding of background conditions.

Paper II. Subjectivism and the Framework of Constitutive Grounds

Paper II deals with a very Lundian problem: How to understand the framework of constitutive grounds and how it may allow us to formulate a version of subjectivism where every value-bearer or value-maker needs to be an attitude.

It should perhaps be noted that neither Andrés Garcia nor I subscribe to subjectivism. The point of the paper is to find a way to make sense of the disagreement between subjectivism and objectivism.

In the first part of the paper we describe the framework of constitutive grounds and how philosophers have applied the framework of constitutive grounds to make sense of the disagreement between subjectivism and objectivism. The framework understands the two theories as being involved in a disagreement about the extent to which value is determined by attitudes. Although the framework affords us some useful observations about how this should be interpreted, the question how value can be determined by attitudes in the first place is left largely unanswered.

The second part of the paper aims to explore the benefits of a positive interpretation which aims to address this oversight and make the framework more informative. We argue for an interpretation that is inspired by recent work of Schroeder (2007)and Sobel (2016).

Our interpretation claims that the relevant sense in which value can be determined by attitudes is discovered by seeing how facts can be endowed with the normative property of being a reason. We argue that this interpretation allows us to significantly deepen our understanding of the disagreement between subjectivism and objectivism.

As can be seen as a general thread throughout most of the papers, the paper relies on a Reasons-First approach or at least that the FA analysis of value is largely correct. On our interpretation of the framework of constitutive grounds turns out to be, or at least share a lot of the features of metaphysical grounding (see pp. 65-68).

So on our interpretation of the framework of constitutive grounds of value we determine that the disagreement between subjectivism and objectivism about value comes down to whether it is always the case that a reason has the property of being a reason partly in virtue of some attitude.

Since the publication of this paper it has come to our attention that one of the originators of the framework of constitutive grounds, Wlodek Rabinowicz, favors a different interpretation than ours. In personal communication he has suggested that a better analogy for how to understand the role of attitudes for a subjectivism

given the framework of constitutive grounds is that attitudes play the same role as God does in the cosmological argument for God's existence.[118]

Paper III. A Fair Reading of 'Ought implies Can'

In this paper I argue for a specific reading of 'can' in the dictum 'ought implies can'. 'Ought implies can' (OIC) stipulates that it could never be the case that you ought to do something that you cannot do.

In the first part of the paper I establish a methodology for how to compare different interpretations of 'can' in OIC. The methodology consists in testing whether the different readings manages to rule out unfair moral requirement ascriptions.

In the second part I argue that most readings of 'can' in the debate fail to rule out unfair moral requirements.

In the third part I propose a new reading of '*S* can *φ*', in a sense that is relevant to OIC. On my proposal '*S* can *φ*' if and only if it is true that if *S* tried her best to *φ*, *S* would *φ*. I argue that it is only on my reading that we are able to retain a strong connection between what we ought to do and what we are blameworthy for failing to do. I, however, also open up for the possibility that the connection between what we ought to do and what we are blameworthy for failing to do might not be as strong as we might have expected and that it might be that my analysis of 'can' in the dictum OIC is false, and that it instead is only true for the closely related dictum 'blameworthy implies can'.

In the fourth part of the paper I outline what it means to try do something, and more specifically what it means to try *one's best* to do something. This endeavor turns out to be harder that what one might expect.

It should be noted that my philosophical views have evolved since writing the paper and that I am now (as is apparent from my writing on deontic buck-passing in the chapter "On Buck-Passing") more friendly towards the suggestion that my reading of 'can' in the paper concern moral obligations rather than moral oughts.

---

[118] The version of the cosmological argument he has in mind is roughly that if everything has a cause, then there must be a first un-caused cause that is essentially different from other causes.

Paper IV. Normative Transmission and Necessary Means

In this paper, which was the first paper that I wrote as a part of my thesis, I investigate the idea of normative transmission, i.e. that normativity can 'transmit' from one action to another. I defend the two principles *Ought Necessity* and *Strong Necessity* against some recent objections from Broome (2013) and Kolodny (2018).

Ought Necessity claims that if an agent ought to φ and ψing is a necessary means of φing then, and because of that, the agent ought to ψ. *Strong Necessity* likewise claims that if an agent has a reason with a certain degree of strength, *w,* to φ and ψing is a necessary means of φing then, and because of that, the agent has a reason with a degree of strength, *w*, to ψ.

Just as Paper III, this paper operates somewhere in between classical meta-ethics and normative ethics, in what might be described as structural normativity. In the first three parts of the paper I outline the principles and argue that denying these principles has consequences for deontic logic that *ceteris paribus* it would be best to avoid. I also make it explicit that the paper is not to be situated in the more familiar discussion about instrumental rationality, but is a standalone problem independent on one's stance within the debate of instrumental rationality.

The fourth part of the paper describes the objections from Broome. The objection, roughly, claims that if we know that we will in fact not φ even if we take the necessary means, ψ, then ψing (as far as φing is concerned) is pointless and that in this scenario no ought, or reason, transmit from φing to ψing. I argue that if we in fact will not φ even if we ψ then it is not clear in which way ψing is a necessary means in the first place. Furthermore, if we describe the scenario such that it is no longer the case that we in fact will not φ if we ψ our intuitions about Broome's counter-examples change in favor of Ought Necessity and Strong Necessity.

The fifth part of the paper analyzes Kolodny's objection. Kolodny sets up a counter-example in which it is intuitive that one agent ought to take the necessary means to φ and that it is not the case that another agent ought to take the necessary means to φ, for all intents and purposes, are identical. My response to Kolodny's objection involves a denial of the characterization that they ought to φ in the first place. I argue that they can only try their best to φ, and it is only the case that the first agent ought to try his best to φ but it is not the case that the second agent

ought to try his best to φ. The difference in whether the agents ought to try their best to φ explains why it is the case that the first, but not the latter, agent ought to take the necessary means of trying their best to φ.

In the paper we see the first instance of what I call 'the test' for fair moral requirement that planted the theoretical seed for Paper III.

Paper V. How Valuable is It?

Paper V concerns an underappreciated complexity on the FA analysis of value (see the chapter "On Buck-Passing"). In the literature the analysis has mostly focused on what it is for something to be valuable, but little has been said about *degrees* of value, i.e. what it is to be very valuable or somewhat valuable.

This is not to say that the discussion about the FA analysis and degrees of value does not have history. Historically, G.E Moore (1903) and Franz Brentano (1889/2009) had different views on how the FA analysis should understand degrees of value. The FA analysis roughly says that an object is valuable because it is fitting to (dis)favor the object. The *moorean proposal* is that if something is, say, very valuable then it is very fitting to (dis)favor the object. The *brentanoean proposal* is to say that it is fitting to (dis)favor the object a lot. The difference between the two proposals is whether one has the normative component, i.e. the 'fittingness', or the attitudinal component, i.e. the 'favoring' to account for degrees of value.

In the paper Henrik Andersson and I start by investigating the moorean proposal and how the normative component is best understood. We find out that it is not uncommon for the normative component to be formulated in non-gradable terms, such as *conclusive reason* or *ought*. Such formulations are of course incompatible with the moorean proposal.

In the second part of the paper we describe the brentanoean proposal. One of the main objections against this view is that if 'valuable' means 'worthy of love' then 'more valuable' must mean 'more worthy of love'. Our answer to this is that various other elaborations of the FA analysis to other kinds of values modify the attitudinal component, such as personal value, instrumental value, and final value (for more see pp. 68-69) and that it is only an objection against semantic version of the FA analysis not metaphysical.

In the third part of the paper we discuss hybrid views, i.e. that it is both the attitudinal- and the normative component that accounts for degrees of value. We reject this proposal in favor of the brentanoean proposal but try to meet the moorean proposal half way. We do so by outlining how the normative component may still play an underappreciated role when it comes to the value of an object – the resilience of an object's value. By the original new dimension of value that we call resilience we mean that the normative component might make the object's value more robust against changes in the underlying reasons to favor the object without the object's value being affected.

The fifth part of the paper compares the two proposals with regards to how they fare when it comes to comparative value relations. We find that both proposals are compatible with the standard accounts of value relations but that the brentanoean proposal has a better fit.

The findings from this paper is first and foremost one of the first structural discussions about the issue of degrees of value and some evidence in favor of the brentanoean proposal. It also opens up the possibility for future research in how the normative- and attitudinal component can be made to shed some light on the new dimension of value, resilience, and possibly of other philosophically interesting concepts, such as credence.

References

Anderson, E. (1993). *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.

Andersson, H. (2016). Vagueness and Goodness Simpliciter. *Ratio, 29*(4), 378-394.

Andersson, H. (2017). *How It All Relates*. Lund: Media-Tryck

Arrhenius, G., & Rabinowicz, W. (2015). Value Superiority. In I. Hirose & J. Olson (Eds.), *The Oxford Handbook of Value Theory*: Oxford University Press USA.

Ashwell, L. (2014). The Metaphysics of Desire and Dispositions. *Philosophy Compass, 9*(7), 469-477.

Baier, K. (1958). *The Moral Point of View*. Itacha: Cornell University Press.

Baumgardt, D. (1946). Legendary Quotations and Lack of References. *Journal of the History of Ideas, 7*(1), 99-102. doi:10.2307/2707271

Bedke, M. (2011). Passing the deontic buck. *Oxford Studies in Metaethics: Volume 6, 6*, 128.

Bergström, L. (1977). Utilitarianism and future misstakes. *Theoria, XLIII*(2), 84 - 102.

Berker, S. (forthcoming). A Combinatorial Argument against Practical Reasons for Belief. *Analytic Philosophy*.

Betti, A. (2015). *Against facts*. Cambridge, MA: MIT Press.

Bliss, R. L. (2013). Viciousness and the structure of reality. *Philosophical Studies, 166*(2), 399-418. doi:10.1007/s11098-012-0043-0

Brentano, F. (1889/2009). *The Origin of Our Knowledge of Right and Wrong* (R. M. Chisholm, Trans.).

Broad, C. D. (1930). *Five Types of Ethical Theory*: London Routledge & Kegan Paul.

Broome, J. (1991). *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Wiley-Blackwell.

Broome, J. (1997). Reasons and Motivation. *Aristotelian Society Supplementary Volume, 7*(1), 131-146.

Broome, J. (2004a). Reasons. In R. J. Wallace, P. Pettit, S. Scheffler, & M. Smith (Eds.), *Reason and Value: Themes From the Moral Philosophy of Joseph Raz* (pp. 28 - 55). Oxford: Oxford University Press.

Broome, J. (2004b). *Weighing Lives* (Vol. 116). Oxford: Oxford University Press.

Broome, J. (2005). Have We Reason To Do As Rationality Requires? - A comment on Raz. *Journal of Ethics and Social Philosophy, 1*.

Broome, J. (2007a). Does rationality concist in responding correctly to reasons? *Journal of Moral Philosophy, 4*, 349-374.

Broome, J. (2007b). Wide or narrow scope? *Mind, 116*(462), 359-370.

Broome, J. (2013). *Rationality Through Reasoning*. Hoboken: John Wiley & Sons.

Brown, C. (2014). The composition of reasons. *Synthese, 191*(5).

Brownlee, K. (2010). Reasons and ideals. *Philosophical Studies, 151*(3), 433-444.

Brunero, J. (2013). Reasons as explanations. *Philosophical Studies, 165*(3), 805-824. doi:10.1007/s11098-012-9982-8

Buchak, L. (2014). Belief, credence, and norms. *Philosophical Studies, 169*(2), 1-27.

Buckland, J. (2017). Normative Reasons Qua Facts and the Agent-Neutral/relative Dichotomy: a Response to Rønnow-Rasmussen. *Philosophia, 45*(1), 207-225.

Bykvist, K. (2009). No Good Fit: Why the Fitting Attitude Analysis of Value Fails. *Mind, 118*(469), 1-30.

Carlson, E. (2016). 'Good' in Terms of 'Better'. *Noûs, 50*(1), 213-223.

Chang, R. (2002a). *Making Comparisons Count*. London: Routledge.

Chang, R. (2002b). The possibility of parity. *Ethics, 112*(4), 659-688.

Chang, R. (2004). All things considered. *Philosophical Perspectives, 18*(1), 1-22.

Chang, R. (2013). Grounding practical normativity: going hyrid. *Philosophical Studies, 164*(1), 163-187.

Chisholm, R. M. (1986). *Brentano and Intrinsic Value*. Cambridge: Cambridge University Press.

Chrisman, M. (2012). On the Meaning of 'Ought'. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics, vol. 7* (pp. 304-332). Cambridge: Oxford University Press.

Copp, D. (1997). Defending the principle of alternate possibilities: Blameworthiness and moral responsibility. *Noûs, 31*(4), 441-456.

Copp, D. (2003). 'Ought' Implies 'Can', Blameworthiness, and the Principle of Alternate Possibilities'. In D. Widerker & M. Mckenna (Eds.), *Moral Responsibility and Alternative Possibilities* (pp. 266-299). Aldershot, U.K: Ashgate.

Correia, F. (2010). Grounding and truth-functions. *Logique Et Analyse, 53*(211), 251-279.

Correia, F., & Schnieder, B. (2012). *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press.

Cresswell, M. J. (1975). Hyperintensional logic. *Studia Logica, 34*(1), 25-38. doi:10.1007/bf02314421

Crisp, R. (2000a). Particularizing Particularism. In B. Hooker & M. O. Little (Eds.), *Moral Particularism* (pp. 23-47). Oxford: Oxford University Press.

Crisp, R. (2000b). Value ... And what follows by Joel Kupperman new York: Oxford university press, £25.00. *Philosophy, 75*(3), 452-462.

Crisp, R. (2006). *Reasons and the Good*. Oxford: Oxford University Press.

D'Arms, J., & Jacobson, D. (2000). Sentiment and value. *Ethics, 110*(4), 722-748.

Dancy, J. (1993). *Moral Reasons*. Oxford: Blackwell Publishing.

Dancy, J. (2000). The Particularist's Progress. In B. Hooker & M. O. Little (Eds.), *Moral Particularism* (pp. 130-156). Oxford: Oxford University Press.

Dancy, J. (2003). Are there organic unities? *Ethics, 113*(3), 629-650.

Dancy, J. (2004). *Ethics Without Principles*. Oxford: Oxford University Press.

Dancy, J. (2015). Reasons for Broome. In A. Reisner & I. Hirose (Eds.), *Weighing and Reasoning: A Festshrift for John Broome* (pp. 177-188). Oxford: Oxford University Press.

Danielsson, S., & Olson, J. (2007). Brentano and the Buck-Passers. *Mind, 116*(463), 511-522.

Darwall, S. L. (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.

deRosset, L. (2013). Grounding explanations. *Philosophers' Imprint, 13*, 1-26.

Dietrich, F., & List, C. (2017). What matters and how it matters: A choice-theoretic representation of moral theories. *Philosophical Review, 126*(4), 421-479.

Elson, L. (2014). Heaps and Chains: Is the Chaining Argument for Parity a Sorites? *Ethics, 124*(3), 557-571.

Elson, L. (2017). Incommensurability as vagueness: a burden-shifting argument. *Theoria*(83), 341-363.

Ewing, A. C. (1939). A suggested non-naturalistic analysis of good. *Mind, 48*(189), 1-22.

Ewing, A. C. (1948). *The Definition of Good*. Westport Conn: Hyperion Press.

Falk, W. D. (1986). *Ought, Reasons, and Morality the Collected Papers of W.D. Falk : With a Foreword by Kurt Baier*. Itacha, N.Y.: Cornell University Press.

Fine, K. (2012). Guide to Ground. In F. Correia & B. Schnieder (Eds.), *Metaphysical Grounding* (pp. 37-80). Cambridge: Cambridge University Press.

Fischer, J. M., & Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Fogal, D. (2016). Reasons, Reason, and Context. In E. Lord & B. Maguire (Eds.), *Weighing Reasons*. Oxford: Oxford University Press.

Fox, C., & Feis, G. (2018). 'Ought implies Can' and the law. *Inquiry: An Interdisciplinary Journal of Philosophy, 61*(4), 370-393.

Garcia, A. G. (2018a). *Between values and the World - Studies in second-order value theory (dissertation)*. Lund: Media-Tryck, Lund University.

Garcia, A. G. (2018b). The Finality and Instrumentality of Value in a Way. *Philosophia, 47*(681), 1-12.

Gert, J. (2007). Normative strength and the balance of reasons. *Philosophical Review, 116*(4), 533-562.

Gert, J. (2012). *Normative Bedrock: Response-Dependence, Rationality, and Reasons*. Oxford: Oxford University Press.

Gertken, J., & Kiesewetter, B. (2017). The right and the wrong kind of reasons. *Philosophy Compass, 12*(5), e12412.

Gibbard, A. (1990). *Wise choices, apt feelings : a theory of normative judgement*. Oxford: Clarendon Press.

Gilbert, M. (1990). Walking Together: A Paradigmatic Social Phenomenon. *Midwest Studies In Philosophy, 15*(1), 1-14.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge M.A.: Harvard University Press.

Guindon, B. (2016). Sources, reasons, and requirements. *Philosophical Studies, 173*(5), 1253-1268. doi:10.1007/s11098-015-0544-8

Gustafsson, J. E. (2013). Indeterminacy and the Small-Improvement Argument. *Utilitas, 25*(4), 433-445.

Gustafsson, J. E. (2014). Neither 'Good' in Terms of 'Better' nor 'Better' in Terms of 'Good'. *Noûs, 48*(1), 466-473.

Hare, R. M. (1952). *The Language of Morals*. Oxford: Clarendon Press.

Harsanyi, J. C. (1986). *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

Heuer, U. (2010). Reasons and impossibility. *Philosophical Studies, 147*(2), 235-246. doi:10.1007/s11098-008-9285-2

Heuer, U. (2011). Beyond Wrong Reasons: The Buck-Passing Account of Value. In M. Brady (Ed.), *New Waves in Metaethics* (pp. 166-184). London: Palgrave Macmillan UK.

Hieronymi, P. (2004). The force and fairness of blame. *Philosophical Perspectives, 18*(1), 115-148.

Horty, F. J. (2012). *Reasons as Defaults*. Oxford: Oxford University Press.

Howard-Snyder, F. (1997). The Rejection of Objective Consequentialism. *Utilitas, 9*(02), 241-248. doi:doi:10.1017/S0953820800005306

Jackson, F., Pettit, P., & Smith, M. (2000). Ethical particularism and patterns. In B. Hooker & M. O. Little (Eds.), *Moral Particularism* (pp. 79-99). Oxford: Oxford University Press.

Kavka, G. S. (1983). The Toxin Puzzle. *Analysis, 43*(1), 33-36.

Kearns, S., & Star, D. (2009). Reasons as Evidence. *Oxford Studies in Metaethics, 4*, 215-242.

Kearns, S., & Star, D. (2013). Weighing Reasons. *Journal of Moral Philosophy, 10*(1), 70-86.

Kearns, S., & Star, D. (2015). Weighing Explanations. In A. Reisner & I. Hirose (Eds.), *Weighing and Reasoning: A Festschrift for John Broome*. Oxford: Oxford University Press.

Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.

Kim, J. (1994). Explanatory Knowledge and Metaphysical Dependence. *Philosophical Issues, 5*, 51-69. doi:10.2307/1522873

Kolodny, N. (2005). Why Be Rational? *Mind, 114*(455), 509-563. doi:10.1093/mind/fzi509

Kolodny, N. (2018). Instrumental Reasons. In D. Star (Ed.), *The Oxford Handbook of Reasons and Normativity* (pp. 731-763). Oxford: Oxford University Press.

Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.

Lang, G. (2008). The Right Kind of Solution to the Wrong Kind of Reason Problem. *Utilitas, 20*(4), 472-489. doi:10.1017/S0953820808003282

Lord, E., & Maguire, B. (2016). *Weighing Reasons*. New York, N.Y.: Oxford University Press.

Maguire, B. (2016). The Value-Based Theory of Reasons. *Ergo: An Open Access Journal of Philosophy, 3*, 233-262.

Maguire, B. (2017). There Are No Reasons for Affective Attitudes *Mind, 127*(507), 779-805.

Manne, K. (2014). Internalism about reasons: sad but true? *Philosophical Studies, 167*(1), 89-117.

Mason, E. (2015). Moral ignorance and blameworthiness. *Philosophical Studies, 172*(11), 3037-3057.

McConnell, T. C. (1980). Utilitarianism and Supererogatory Acts. *Ratio, 22*(1), 3-38.

McHugh, C., & Way, J. (2016). Fittingness First. *Ethics, 126*(3), 575-606.

McKeever, S., & Ridge, M. (2005). What does holism have to do with moral particularism? *Ratio, 18*(1), 93-103.

McKeever, S., & Ridge, M. (2006). *Principled Ethics: Generalism as a Regulative Ideal*. Oxford: Oxford University Press.

McKeever, S., & Ridge, M. (2013). Why Holists Should Love Organic Unities. In D. Bakhurst, M. O. Little, & B. Hooker (Eds.), *Thinking About Reasons: Themes From the Philosophy of Jonathan Dancy* (pp. 265-285). Oxford: Oxford University Press.

McKenna, M. (2012). *Conversation & Responsibility*. Oxford: Oxford University Press.

McNamara, P. (2014). "Deontic Logic". In E. N. Zalta (Ed.), *The Standford Encyclopedia of Philosophy* (Winter 2014 ed.): Metaphysics Research Lab, Stanford University.

McNaughton, D., & Rawling, P. (2000). Unprincipled Ethics. In B. Hooker & M. O. Little (Eds.), *Moral Particularism* (pp. 256-275). Oxford: Oxford University Press.

Mill, J. S. (1998). *Utilitarianism*. Oxford: Oxford University Press.

Moore, G. E. (1903). Book Review: The Origin of Knowledge of Right and Wrong. Franz Brentano. *Journal of Ethics, 14*(1), 115-123.

Moore, G. E. (1993/1903). *Principia Ethica*. Cambridge: Cambridge University Press.

Mulligan, K. (1998). From Appropriate Emotions to Values. *The Monist, 81*(1), 161-188.

Nagel, T. (1970). *The Possibility of Altruism*. Oxford: Clarendon Press.

Nagel, T. (1986). *The View From Nowhere*. Oxford: Oxford University Press.

Nolan, D. (2014). Hyperintensional metaphysics. *Philosophical Studies, 171*(1), 149-160. doi:10.1007/s11098-013-0251-2

Olson, J. (2004). Buck-passing and the wrong kind of reasons. *Philosophical Quarterly, 54*(215), 295-300.

Olson, J. (2009). The wrong kind of solution to the wrong kind of reason problem. *Utilitas, 21*(2), 225-232.

Olson, J. (2013). Buck-Passing Accounts. In H. LaFollette (Ed.), *The International Encyclopedia of Ethics*. Oxford: Wiley-Blackwell.

Olson, J. (2014). *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.

Olson, J. (2018). The Metaphysics of Reasons. In D. Star (Ed.), *The Oxford Handbook of Reasons and Normativity* (pp. 255-274). Oxford: Oxford University Press.

Orsi, F. (2013). Fitting Attitudes and Solitary Goods. *Mind, 122*(487), 687-698.

Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.

Parfit, D. (1997). Reasons and motivation. *Aristotelian Society Supplementary Volume, 71*(1), 99-130.

Parfit, D. (2001). Rationality and Reasons. In D. Egonsson, J. Josefsson, B. Petersson, & T. Rønnow-Rasmussen (Eds.), *Exploring Practical Philosophy* (pp. 17-40). Aldershot: Ashgate.

Parfit, D. (2011a). *On What Matters: Vol. 1* (Vol. 1). Oxford: Oxford University Press.

Parfit, D. (2011b). *On What Matters: Vol. 2* (Vol. 2). Oxford: Oxford University Press.

Parfit, D. (2016). Can We Avoid the Repugnant Conclusion? *Theoria, 82*(2), 110-127.

Portmore, D. W. (2003). Position-relative consequentialism, agent-centered options, and supererogation. *Ethics, 113*(2), 303-332.

Rabinowicz, W. (2008). Value relations. *Theoria, 74*(1), 18-49.

Rabinowicz, W. (2012). Value relations revisited. *Economics and Philosophy, 28*(2), 133-164.

Rabinowicz, W. (2017). From values to probabilities. *Synthese, 194*(10), 3901-3929.

Rabinowicz, W., & Rønnow-Rasmussen, T. (2000). A Distinction in Value: Intrinsic and for its Own Sake. *Proceedings of the Aristotelian Society (Hardback), 100*(1), 33-51. doi:doi:10.1111/j.0066-7372.2003.00002.x

Rabinowicz, W., & Rønnow-Rasmussen, T. (2004). The Strike of the Demon: On Fitting Pro-attitudes and Value. *Ethics, 113*(3), 391-423.

Rabinowicz, W., & Rønnow-Rasmussen, T. (2006). Buck-passing and the right kind of reasons. *Philosophical Quarterly, 56*(222), 114-120.

Raven, M. J. (2013). Is ground a strict partial order? *American Philosophical Quarterly, 50*(2), 191-199.

Raz, J. (1986). *The Morality of Freedom*. Oxford: Oxford University Press.

Raz, J. (2000). The Truth in Particularism. In B. Hooker & M. O. Little (Eds.), *Moral Particularism* (pp. 48-78). Oxford: Oxford University Press.

Raz, J. (2006). Review: The Trouble with Particularism (Dancy's Version). *Mind, 115*(457), 9-120.

Raz, J. (2011). *From Normativity to Responsibility*. Oxford: Oxford University Press.

Reisner, A. (2009a). Abandoning the buck passing analysis of final value. *Ethical Theory and Moral Practice, 12*(4), 379-395.

Reisner, A. (2009b). The possibility of pragmatic reasons for belief and the wrong kind of reasons problem. *Philosophical Studies, 145*(2), 257-272.

Reisner, A. (2015a). Fittingness, Value and Trans-World Attitudes. *Philosophical Quarterly, 65*(260), 464-485.

Reisner, A. (2015b). Normative Conflicts and the Structure of Normativity. In I. Hirose & A. Reisner (Eds.), *Weighing and Reasoning: Themes from the Work of John Broome*. Oxford: Oxford University Press.

Reisner, A. (2018a). Pragmatic Reasons for Belief. In D. Star (Ed.), *The Oxford Handbook of Reasons and Normativity* (pp. 705-728). Oxford: Oxford University Press.

Reisner, A. (2018b). Two Thesis about the Distinctness of Practical and Theoretical Normativity. In C. McHugh, J. Way, & D. Whiting (Eds.), *Normativity: Epistemic and Practical* (pp. 221-240). Oxford: Oxford University Press.

Robertson, S. (2008). Not So Enticing Reasons. *Ethical Theory and Moral Practice, 11*(3), 263-277. doi:10.1007/s10677-007-9091-5

Rodrigues-Pereyra, G. (2015). Grounding is not a Strict Order. *Journal of the American Philosophical Association, 1*(3), 517-534.

Ross, W. D. (1963/1939). *Foundations of Ethics*. Oxford: Oxford University Press.

Ross, W. D. (2002/1930). *The Right and the Good* (P. Stratton-Lake Ed.). Oxford: Oxford University Press.

Rossi, B. (2016). Moral Excuses and Blame-Based Theories of Moral Wrongness. *Southwest Philosophy Review, 32*(1), 153-165.

Rowland, R. (2017). Reasons or Fittingness First? *Ethics, 128*(1), 212-229.

Rønnow-Rasmussen, T. (2009). Normative reasons and the agent-neutral/relative dichotomy. *Philosophia, 37*(2), 227-243.

Rønnow-Rasmussen, T. (2011). *Personal Value*. Oxford: Oxford University Press.

Sampson, E. (2015). Against Scanlon's Theory of the Strength of Practical Reasons. *Journal of Ethics and Social Philosophy, 111*(11), 631 - 640.

Scanlon, T. M. (1998). *What We Owe to Each other*. Cambridge: Belknap.

Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford: Oxford University Press.

Schaffer, J. (2009). On what grounds what. In D. Manley, D. J. Chalmers, & R. Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (pp. 347-383). Oxford: Oxford University Press.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies, 173*(1), 49-100. doi:10.1007/s11098-014-0438-1

Schnieder, B. (2016). In defence of a logic for 'because'. *Journal of Applied Non-Classical Logics, 26*(2), 160-171.

Schroeder, M. (2007). *Slaves of Passion*. Oxford: Oxford University Press.

Schroeder, M. (2010). Value and the right kind of reason. *Oxford Studies in Metaethics, 5*, 25-55.

Schroeder, M. (2012). The Ubiquity of State-Given Reasons. *Ethics, 122*(3), 457-488.

Setiya, K. (2014). What is a Reason to Act? *Philosophical Studies, 167*(2), 221-235.

Shafer-Landau, R. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.

Shah, N. (2006). A new argument for evidentialism. *Philosophical Quarterly, 56*(225), 481-498.

Sharadin, N. P. (2016). Nothing but the Evidential Considerations? *Australasian Journal of Philosophy, 94*(2), 343-361.

Shoemaker, D. (2011). Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics, 121*(3), 602-632.

Sidgwick, H. (1981/1874). *The Methods of Ethics*. Kndianapolis: Hackett.

Sinnott-Armstrong, W. (1984). `Ought' conversationally implies `can'. *Philosophical Review, 93*(2), 249-261.

Skorupski, J. (2010a). *The Domain of Reasons*. Oxford: Oxford University Press.

Skorupski, J. (2010b). Sentimentalism: Its Scope and Limits. *Ethical Theory and Moral Practice, 13*(2), 125-136.

Skorupski, J. (2012). The triplism of practical reason. *Ratio, 25*(2), 127-147.

Smid, J. (2017). What does "nothing over and above its parts" actually mean? *Philosophy Compass, 12*(1), e12391.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Sobel, D. (2016). *From Valuing to Value: Towards a Defense of Subjectivism*. Oxford: Oxford University Press.

Stalnaker, R. (1984). *Inquiry*. Cambridge M.A.: MIT Press.

Stern, R. (2015). Does 'Ought' Imply 'Can'? And Did Kant Think It Does? *Utilitas, 16*(1), 42-61. doi:10.1093/acprof:oso/9780198722298.003.0007

Stratton-Lake, P. (2002). *Introduction: The Right and the Good*. Oxford: Clarendon Press.

Stratton-Lake, P. (2005). How to Deal with Evil Demons: Comment on Rabinowicz and Rønnow-Rasmussen. *Ethics, 115*(4), 788-798. doi:10.1086/430491

Strawson, P. F. (1962). Freedom and resentment. In G. Watson (Ed.), *Proceedings of the British Academy, Volume 48: 1962* (Vol. 48, pp. 1-25). Oxford: Oxford University Press.

Streumer, B. (2007). Reasons and impossibility. *Philosophical Studies, 136*(3), 351-384.

Streumer, B. (2010). Reasons, impossibility and efficient steps: reply to Heuer. *Philosophical Studies, 151*(1), 79-86.

Streumer, B. (2018). I Ought to Reply, So I Can. *Philosophia*. doi:10.1007/s11406-018-0042-2

Sylvan, K. (2015). What apparent reasons appear to be. *Philosophical Studies, 172*(3), 587-606. doi:10.1007/s11098-014-0320-1

Temkin, L. S. (2012). *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford: Oxford University Press.

Toulmin, S. E. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Trogdon, K. (2013). An Introduction to Grounding. In M. Hoeltje, B. Schnieder, & A. Steinberg (Eds.), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence (Basic Philosophical Concepts)* (pp. 97-122). Münich: Philosophia Verlag.

Trogdon, K. (Ms). Grounding and Explanation.

Urmson, J. O. (1958/1969). Saints and heroes. In J. Feinberg (Ed.), *Moral Concepts*. Oxford: Oxford University Press.

Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

Way, J. (2012). Transmission and the Wrong Kind of Reason. *Ethics, 122*(3), 489-515.

Way, J. (2017). Reasons as Premises of Good Reasoning. *Pacific Philosophical Quarterly, 98*(2), 251-270.

Wedgwood, R. (2013). Rational 'ought' implies 'can'. *Philosophical Issues, 23*(1), 70-92.

Vessel, J.-P. (2010). Supererogation for utilitarianism. *American Philosophical Quarterly, 47*(4), 299-319.

Wiggins, D. (1987). *Needs, Values, Truth: Essays in the Philosophy of Value*. Oxford: Basil Blackwell.

Williams, B. (1981a). Internal and External Reasons *Moral Luck* (pp. 101-113). Cambridge: Cambridge University Press.

Williams, B. (1981b). *Ought* and moral obligation *Moral Luck* (pp. 114-123). Cambridge: Cambridge University Press.

Wolf, S. (2015). *The Variety of Values: Essays on Morality, Meaning, and Love*. Oxford: Oxford University Press.

Vranas, P. B. M. (2007). I Ought, Therefore I Can. *Philosophical Studies, 136*(2), 167 - 216.

Wygoda Cohen, S. (2019). Not All Partial Grounds Partly Ground: Some Useful Distinctions in the Theory of Grounding. *Philosophy and Phenomenological Research*. doi:10.1111/phpr.12524

Yamada, M. (2010). A new argument for evidentialism? *Philosophia, 38*(2), 399-404.

Zimmerman, M. J. (2001). *The Nature of Intrinsic Value*. Lanham, Maryland: Rowman & Littlefield.

Reasons and Normativity

Normative reasons concern us all. We constantly make decisions, believe in various propositions, and feel certain emotions. Hopefully, we do, feel and believe what we ought to do, feel, and believe – but how does it all come together? *Reasons and Normativity* is a thesis that consists of five stand-alone papers and an extended introduction that tries to deepen our understanding of how reasons and normativity works.

Lund University
Faculties of Humanities and Theology
Department of Philosophy