Between Values and the World

Studies in second-order value theory

Garcia, Andres

2018

Link to publication

Citation for published version (APA):
Between Values and the World
Studies in second-order value theory

ANDRÉS G. GARCIA
DEPARTMENT OF PHILOSOPHY | FACULTY OF HUMANITIES AND THEOLOGY | LUND UNIVERSITY
Between Values and the World

Value is an inescapable part of the human experience and what life must be like for a conscious and feeling person. Philosophical questions about value are therefore naturally invited: What sort of thing would value be if it were part of the furniture of the world? How should we understand the relations that value is thought to stand in to other things? In a broad sense, these are formal questions calling for philosophical studies into the understanding of value notions. The following work represents one such study. Its main concern lies specifically with the notion of final value, the idea that things can be good or bad for their own sakes.
Between Values and the World

Studies in Second-Order Value Theory

Andrés G. Garcia

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Prof. Toni Rønnow-Rasmussen, Prof. Włodek Rabinowicz
Faculty opponent: Prof. Michael J. Zimmerman

To be publicly defended, with due permission of the Faculties of Humanities & Theology of Lund University, at 09.00 on the 26th of May 2018.
Value is an inescapable part of the human experience and what life must be like for a conscious and feeling person. Philosophical questions about value are therefore naturally invited: What sort of thing would value be if it were part of the furniture of the world? How should we understand the relations that value is thought to stand in to other things? In a broad sense, these are formal questions calling for philosophical studies into the understanding of value notions. The following work represents one such study. Its main concern lies specifically with the notion of final value, the idea that things can be good or bad for their own sakes.

Key words: Final value, intrinsic value, extrinsic value, dependence, resultance
Between Values and the World

Studies in Second-Order Value Theory

Andrés G. Garcia
Acknowledgments

My first thanks go out to my supervisors and philosophical role models, professors Toni Rønnow-Rasmussen and Wlodek Rabinowicz. It was their criticism of my dearly held assumptions as an undergraduate that eventually made me want to become a philosopher. If I should prove unsuccessful in this endeavor, then it will be entirely despite of all the help, encouragement, and patient guidance that they have afforded me over the years. For similar reasons my gratitude is also owed to Fritz-Anton Fritzson, Jakob G. Werkmäster, Henrik Andersson, Frits Gåvertsson, Aron Vallinder, Simon Rosenqvist, Joel Parthemore, Jeroen Smid, Eric Brandstedt, Oscar Ralsmark, Olof Leffler, Morten Dahlback, Christoffer Ivarsson, Gloria Mähringer, Seyyed M. Eslami, Sarah Marie Babbitt, Giuseppe Capone, Demeng Kong, Emma Bäckström, Catherine Felix, Jared Parmer, Björn Petersson, Dan Egonsson, Jonathan Dancy, Michael J. Zimmerman, Erik Olsson, Marcus Agnafors, Andras Szigeti, Jens Johansson, Robin Stenwall, David Alm, Erik Olsson, Stefan Schubert, John Skorupski, Astrid Byrman, Jacob Olofsson, Robin Linderborg, Sara Rahm, Ingvar Johansson, Martin Jönsson, Anna-Sofia Maurin, Christian Tallhamn, Tomas Kronvall, Andrew Reisner, Caj Strandberg, Jonas Josefsson, Anna Östberg, Staffan Angere, Ylva von Gerber, Emma Johansson, Lisa Miracchi, George Masterton, Hanna Fange, Anna Cagnan Enhörning, Kim Andersen, Ben Bramble, Tomas Persson, Tobias Hanson Wahlberg, Lisa Erickson Norberg, Joakim Orrbén, Karl Erickson, Ada Knitter, Johan Nylund, Anna-Sofia Maurin, Paula Quinon, Joakim Almqvist, Daniel Almqvist, Robin Almqvist, Monica Almqvist, Olle Olsson, Karin Björnberg, Anders Björnberg, Anette Björnberg, Staffan Björnberg, Fredrik Henriksson, Berit Braun, Johanna Bäckström, Ellinor
Stråhed, Shirin Golsirat, Madeleine Hedqvist, Jenny Axelsson, Rosanna Ljungren, Emma Johansson, Isabelle Barrett, Jonatan Gyllenör, Moa Gyllenör, Hannes Gyllenör, Per Gyllenör, Mia Gyllenör, Helene Almqvist, José Antonio García Rivero, Lars-Göran Andersson, Ines de los Reyes Rodríguez Macías, Barbro Björnberg, Janne Karlsson, Margit Björnberg, José García García, and Antonia Maria del Carmen Rivero Garcia. Last, but certainly not least, I would like to thank Debora Gyllenör, for loving me and for being so lovable. I can think of nothing that better illustrates the central theme of this book than her. She deserves to have far better works dedicated to her, but this one will have to do for now.
Contents

CHAPTER 1 INTRODUCTION .......................................................................................... 1
  FOUNDATIONS & RESTRICTIONS.............................................................................. 7
    On Normative Explanations.................................................................................. 9
    On the Fitting-Attitudes Analysis...................................................................... 19
  A BRIEF RESTATEMENT OF THE PROJECT.......................................................... 36

CHAPTER 2 CHALLENGES FROM ATTRIBUTIVITY............................................. 37
  THE CASE AGAINST FINAL VALUE ...................................................................... 40
  DESCRIPTIVISM ABOUT VALUE IN A WAY ............................................................ 52
  THE FITTING-ATTITUDES ANALYSIS OF VALUE IN A WAY .................................... 57
    Toward a Conditional Approach ........................................................................ 59
    Toward a Nonconditional Approach .................................................................. 67
  CONCLUDING REMARKS ...................................................................................... 72

CHAPTER 3 CHALLENGES FROM DEPENDENCE .......................................... 75
  ANALYZING EVALUATIVE DEPENDENCE............................................................. 78
  ON THE ENTAILMENT DILEMMA ........................................................................ 88
    On Property-Preserving Predicates .................................................................... 93
    From Necessary Coextension to Identity .......................................................... 101
  REMARKS ON THE CONCEPT OF GROUNDING .................................................. 106
    Skepticism about Grounding ........................................................................... 109
  CONCLUDING REMARKS .................................................................................... 119

CHAPTER 4 ON BACKGROUND CONDITIONS................................................. 121
  PREPARING THE GROUND ................................................................................... 125
  THE DEFLATIONARY VIEW .................................................................................. 129
CHAPTER 5 ON CONSTITUTIVE GROUNDS .................................................. 159
INTERPRETING THE FRAMEWORK ......................................................... 164
On Evaluative Projectivism ................................................................. 165
On the Chess Analogy ............................................................... 171
Toward a Hyperintensional Interpretation ........................................... 176
CONCLUDING REMARKS ................................................................. 184

CHAPTER 6 ON EVALUATIVE HOLISM ................................................. 187
REMARKS ON THE RIGHT KIND OF PARTS AND VALUES ...................... 189
The Substantive Failure of Evaluative Atomism ..................................... 193
Hints of Holistic Behaviors ................................................................. 198
On the Mechanisms of Evaluative Holism ............................................ 204
A Transmission Principle for Parts and Wholes .................................. 210
Unfitting Ambivalences and Normative Asymmetries ......................... 219
On the Prosperity of the Sinful and the Possibility of Hybrid Views ........ 225
CONCLUDING REMARKS ................................................................. 228

CHAPTER 7 CLOSING REMARKS .......................................................... 231
APPENDIX ON CIRCULAR ANALYSES OF VALUE ................................. 237
The Intuitive Case for Circular Benignity ......................................... 239
Informativity and the Ban on Inferential Circularity ............................ 245
Wrong-Kind Reasons and the Virtues of Circularity .............................. 250
REFERENCES ................................................................................... 255
Chapter 1

Introduction

As human beings, we cannot help but judge things to be good or bad. It seems to happen automatically and without the consent of our deliberate reason. Sometimes we even have experiences of the good and bad. We see the admirability in an act of kindness, just as we see beauty in the setting sun.\(^1\) Value is therefore an inescapable part of the human condition and what life must be like for a conscious and feeling person. It is not surprising, then, that studies of the good and bad occur in many areas of research, characterized by a wide variety of aims and methodologies.\(^2\) The work you are about to read could be said to constitute a distinctly philosophical study of value, but this would be too vague a statement. For it turns out that even a study of this type can proceed

\(^{1}\) The example is borrowed from David McNaughton: “I do not see an expanse of coloured cloud, which is not itself seen as beautiful, and then experience a thrill of pleasure to which I give the name beauty. The beauty of the sunset is woven into the fabric of my experience of it. I see the sunset as beautiful” (1988, p. 56). It should perhaps be noted that we can agree with this irrespective of what we think about the ontological status of value. Admitting that we have evaluative experiences does not force us to understand value as a genuine property or relation in the world. The following work aims to be neutral about ontological questions of this sort.

\(^{2}\) In a broad sense, empirical studies of value take place in psychology, anthropology, gender studies, sociology, economics, political science, and virtually any other field that focuses on the minds and social lives of human beings. The goals of these empirical studies vary quite a bit, but they are often about establishing what value judgments people hold, why people hold the value judgments that they have, how value judgments tend to develop over time, and what effect value judgments have on both an individual and societal scale. While empirical studies of value sometimes take philosophical issues into account, their main aims are not explicitly philosophical.
in several distinct ways. To start with, we should keep in mind that there is a
 distinction in value theory between first-order questions and second-order
 questions. While the former questions tend to be about the substantive issue of
 what objects have value and why, the latter questions are concerned with the
 more abstract issue of what value is and how it should be understood. What
 follows are some very basic examples that are meant to illustrate the differences
 between the two categories of inquiry:

First-Order Questions: What objects are good? What objects are bad?
What makes happiness desirable? What makes a person respectable?
What does the good life consist in? What objects should be promoted
as ends? Does symmetry have a greater contribution to beauty than
proportionality? Is individual freedom ultimately better than equality?

Second-Order Questions: Is value a genuine property or relation? How
do we go about justifying value judgments? What are the semantic
features of value statements? What connections do value judgments
have to motivation? What value concepts are there? What features do
value concepts have? How can value concepts be analyzed?

Among the observations that can perhaps be gleaned from these examples is that
whereas answering first-order questions will require us to endorse actual value
judgments, answering second-order questions will not. Instead, making
progress on second-order questions will force us to endorse more abstract
judgments, either about value itself or about the nature of evaluation. The work
you are about to read finds itself comfortably situated within the domain of
second-order questions, or simply the second-order domain. It is specifically
concerned with illuminating the dependencies and relations that value stands in
vis-à-vis other things. Dealing with this issue means that several different

3 We might be required to endorse value judgments about the relative merits of different
second-order theories, but this, of course, is not quite the same thing. To endorse one of those
second-order theories is not the same as endorsing any first-order value judgment.

4 Second-order questions about value have been asked since the dawn of philosophy, but the
person who is most responsible for making it into its own delineated field is arguably G. E.
Moore. Contemporary discussions about the nature of value are often inspired by, or
formulated in reaction to, the arguments found in his seminal Principia Ethica (1993/1903).

5 It has become standard practice in many areas of research to attempt to come to grips with
unpalatably mysterious phenomena by clarifying how they relate to other things. Value is no
exception. These considerations are not only of a conceptual interest but have also been
thought to be of relevance to how we understand the ontological status of value. J. L. Mackie
second-order questions will have to be considered, including what value concepts there are, what features they have, and how they should all be analyzed. Of course, this does not mean that first-order questions will be left completely by the wayside. It just means that whenever first-order questions are addressed, and any first-order judgments are expressed, this is mainly so that they can help me highlight the more abstract commitments that seem to be involved in our evaluative talk and practices.

To make the following endeavors more manageable, most of my attention will here be focused on the concept of *final value*. This concept is meant to pick out the type of value that accrues to objects that are good or bad for their own sakes. Although it can be distinguished from a great deal of other value concepts, final value is traditionally contrasted with the concept of *instrumental value*. This concept picks out the type of value that accrues to objects that are good or bad for the sake of their effects. One reason that the concept of final value should be of special interest to us is that it occupies a very central place in our evaluative talk and practices. This becomes especially clear whenever it is couched in the more familiar terms of ends and means. For example, the wellbeing of our loved ones is something that a lot of people would prefer to treat as an end in itself. By contrast, money is something that a lot of people are likely treat as a means toward other things. Indeed, a person who was known to treat money as an end in itself would risk being regarded with bafflement by other people (Orsi, 2015, pp. 25–26).

If this pattern of value judgments is as common as it appears to be, this would suggest that people tend to judge the wellbeing of our loved ones as having final value, while money is only thought to have instrumental value. Most cases are of course likely to be more divisive and harder to predict than these. For example, some people might treat knowledge and the development of artistic talent as worthy of being promoted as ends in themselves, while others only treat these things as means toward other things. The disagreement in question is an interesting first-order disagreement about what has final value and why. Yet

(1977) famously rejects the notion that value should be regarded as a genuine property and relation in the world, at least partly because of the difficulties involved with getting a good grasp of its dependencies and relations. Variations on this line of reasoning, to which we shall return, have also been developed by other philosophers, including Blackburn (1971, 1985).

6 The term ‘object’ should be understood in a wide sense, as something that could constitute concrete objects, states of affairs, facts, relations, properties, structures, events, and so on.

7 That is unless, perhaps, they happen to be coin collectors, in which case we might give them a free pass. Examples like these will be discussed in an upcoming subsection dealing with the possibility of *extrinsic final value*. 

---

(1977) famously rejects the notion that value should be regarded as a genuine property and relation in the world, at least partly because of the difficulties involved with getting a good grasp of its dependencies and relations. Variations on this line of reasoning, to which we shall return, have also been developed by other philosophers, including Blackburn (1971, 1985).

6 The term ‘object’ should be understood in a wide sense, as something that could constitute concrete objects, states of affairs, facts, relations, properties, structures, events, and so on.

7 That is unless, perhaps, they happen to be coin collectors, in which case we might give them a free pass. Examples like these will be discussed in an upcoming subsection dealing with the possibility of *extrinsic final value*. 

---

---
couching the disagreement in terms of ends and means reminds us of something important about the abstract and second-order nature of value. Among other things, it shows that the degree to which we are in a genuine first-order disagreement in the first place depends upon the type of value that we have in mind. For although people may agree that objects like knowledge and the development of artistic talent have value, not everyone will agree that these objects are good or bad for their own sakes. This brings me to the second reason why the concept of final value should be of special interest to us. Besides occupying a principal place in our evaluative talk and practices, the concept seems significant from the perspective of moral philosophy.

The assumption here is that ordinary disagreements about what is good and bad, as well as systematic attempts to formulate ethical theories, are often grounded in notions about what objects have final value. That this assumption has been and continues to be a common one can be given some measure of support by looking at examples from the history of moral philosophy (Orsi, 2015, pp. 27–31; Rønnow-Rasmussen & Zimmerman, 2005, pp. 14–15). Consequentialism famously states that we ought to act in such a way that we maximize the amount of utility in the world. The theory supplements this deontic claim with an evaluative one, which says that what is ultimately good for its own sake is something like happiness, preference satisfaction, or the experience of pleasure. Kantianism, on the other hand, insists that we ought to act in accordance with the maxims we could willingly see become universal laws, adhered to by any agent with the capacity of reason. The theory weds this deontic claim to an evaluative one, which states that what is ultimately good for its own sake is the will to act in accordance with duty. Other examples can be cited, but these are sufficient to hint at the theoretical and practical usefulness of the concept of final value.

---

8 Utilitarianism also claims that what is bad for its own sake is the absence of the above-mentioned things, or something like unhappiness, preference frustration, and the experience of pain. Kantianism insists that what is bad for its own sake is the sullied will, which, roughly, compels agents to act in a way that is contrary to their moral duty.

9 Of course, just because the concept of final value is potentially important, from the perspective both of everyday lives and moral philosophy, we should take care not to assume that all cases of final value must be very weighty from a normative standpoint. Besides things like happiness, pleasure, preference satisfaction, and the will to act in accordance with duty, final value can also be ascribed to categories of objects of a more mundane nature, like rare coins, historically significant pens, and an untouched wilderness. The point, then, is that final value is meant to be distinct from the concept of being the most valuable and even the greatest value. Not all ends are created equal. In addition, it will be pointed out soon that the reasons for attitudes that
The structure of this work is relatively modular, meaning that each of its six chapters could be read largely independently of the others. However, the book can also be understood as consisting of two main parts: The first half of the book, consisting of chapter two and chapter three, is mainly defensive in character. It is meant to show that the concept of final value and one of its most central dependencies are meaningful subjects of philosophical inquiry. More specifically, chapter two focuses on an influential objection which draws its inspiration from the discussions of Peter Geach (1956) and Judith Jarvis Thomson (2008). The objection states that the concept of final value is incoherent, since everything that has value must be good or bad relative to kinds or for something. Besides showing that the objection is weaker than often thought, the chapter offers an argument to the effect that final value is a rather fundamental concept in value theory. What this means is that many other value concepts that are represented in our evaluative talk and practices, including those of having value in the senses just mentioned, can come in the final variety. Consequently, one of the more constructive upshots of this discussion is an elucidation of the relation between final value and those other value concepts.

Chapter three focuses on a somewhat similar challenge, not against the concept of final value as such, but against one of its more central dependencies. These dependencies are assumed to have a close connection to a rather thick notion of explanation and priority. To say that final value depends upon a factor, or that it results from it, is to say that it is explained by it. The chapter aims to explicate the most central dependencies that final value stands in and clarify what they commit us to. Some philosophers, like Thomas Hofweber (2009) and Christopher Daly (2012), have recently insisted that aspects of these dependencies are problematic. Part of the reason is that the dependencies involve a notion of priority that cannot be analyzed, and which lacks cognitive content. Other philosophers, like Jessica Wilson (2014) and Kathrin Koslicki (2015), argue that the dependencies in question are theoretically dispensable. If these objections are right, it may seem to spell disaster for the accounts to be developed and defended here, since some of those accounts are couched in terms of relations like these. The chapter presents an argument to the effect that final value appears to correspond to need not be moral or indeed strong enough to amount to any other sort of ought or rational requirement.

---

10 Other philosophers that have made points along these lines include Foot (1985, pp. 196–209) and Hursthouse (1999, pp. 195–206). However, the semantic and conceptual attack on final value is to my mind most clearly and forcefully expressed in the previously mentioned works by Geach and Thomson. For this reason, their discussions will be used as points of departure.
dependencies are meaningful and significant tools in philosophical theorizing, although there may certainly be legitimate doubts about their ontological status.

Once we are done with the first half of this book, we move on to its second part which is meant to be more constructive in character. In chapter four we begin to take a closer look at some recent attempts by philosophers to expand upon our understanding of how final value can be affected by the context of its occurrence. The first attempt to be discussed is primarily inspired by the writings of Jonathan Dancy (1993, 2004). In these writings, he hopes to show that final value can be sensitive to background conditions that help explain why objects end up being good or bad for their own sakes, without thereby being among the factors in virtue of which, or as a result of which, the objects have final value. The purpose of this chapter is to offer an explication of background conditions that serve as an alternative to some other common accounts. Unlike some of those accounts, the explication developed and defended manages to be both faithful and simple. It is faithful in the sense of managing to accommodate the first-order intuitions that tend to draw our attention to the concept of background conditions in the first place. It is simple in the sense of being theoretically parsimonious, since it avoids having to introduce new types of relation alongside the dependencies mentioned above. Roughly, the explication understands the background conditions for final value in terms of the factors in virtue of which, or as a result of which, facts become reasons for attitudes.

The second attempt to expand upon our understanding of how final value can be affected by context, discussed throughout chapter five, involves the framework of constitutive grounds. The framework was originally developed by Wlodek Rabinowicz (Rabinowicz & Österberg, 1996) to make sense of preferentialism, the first-order view that connects value to the satisfaction of preferences. The framework was later elaborated upon by Johan Brännmark (2002), Toni Rønnow-Rasmussen (2003), and Fritz-Anton Fritzson (2014). The question is whether the framework can also be used to illuminate the theory of subjectivism and its claim that final value is determined by the presence of attitudes. The argument will here be made that while the framework of constitutive grounds may seem promising, it needs a positive interpretation if it is to be as informative we want it to be. The interpretation which will here be developed and defended is inspired by some of the recent writings of Mark Schroeder (2007), David Sobel (2016) and others, and largely mirrors the previously mentioned explication of background conditions. Roughly, the interpretation states that constitutive grounds of value should also be understood in terms of the things in virtue of which, or as a result of which, facts become reasons for attitudes. While subjectivism claims that the presence of attitudes is
responsible for providing facts with their normative status, the competing theory of objectivism should be taken to deny that this is necessarily so.

In the sixth chapter of the book, we shall move on and take a closer look at issues having to do with the structure and behavior of final value. More specifically, the discussion will focus on the relation between the final value of complex wholes and the value of their respective parts. We shall be specifically motivated to consider two different versions of evaluative holism: The first version, inspired by G. E. Moore’s (1993/1903, p. 79) principle of organic unities, states that the final value of complex wholes need not be equal or even proportional to the sum value of their respective parts. The second version of holism, on the other hand, states that the final value of an object can change or be lost altogether depending upon the context of its occurrence. Here the concepts of background conditions and constitutive grounds will once again gain relevance. Although much of the chapter will be dedicated to figuring out how we should understand each version of evaluative holism, an attempt will also be made to defend the former by answering some recent objections with which it has been faced. The objections were formulated by Jonathan Dancy (2003), Jonas Olson (2004), and Michael J. Zimmerman (2014) and were meant to throw doubt over the possibility of complex wholes that are good for their own sakes and yet consist of only bad or worthless parts. Finally, the book ends with an appendix where I defend an assumption regarding the analysis of final value that I will use throughout the book. I will say more about this assumption in the next subsection.

Foundations & Restrictions

With this brief overview in hand, I wish to say something now about the general framework that is applied in the discussions just outlined. As endeavors are made to clarify the relations and dependencies of final value, the following work will be informed by certain key assumptions. When taken together, these

---

11 This version of holism is often discussed in conjunction with particularism, which states that the reasons we have cannot be captured by general normative principles (e.g. McDowell, 1979; McNaughton, 1988; McNaughton & Rawling, 2000; Dancy, 1993, 2000, 2004; Little, 2000). In fact, holism has been assumed to give direct support to particularism, but this assumption has since been the subject of doubt among philosophers (e.g. Lippert-Rasmussen, 1999; McKeever & Ridge, 2005).
assumptions constitute the explanatory framework which will be applied in the upcoming chapters. It is necessary, therefore, that these be carefully explicated and at least partly motivated. To this end, some of the assumptions will be discussed here in their order of contentiousness, going from what is likely to be the least controversial to what is likely to be the most controversial. One of the views that will be discussed here is that normative properties generally and final value specifically are not free-floating entities. This is to say that when an object is good or bad for its own sake, we are committed to there being other factors that explain why this is so. The first subsection to follow gives a rough characterization of this intuition, with the intent of clarifying the minimal understanding that is necessary for the explanatory framework to work and be understood. In so doing, it will feature an attempt on my part to distinguish the dependence of final value from the relation of evaluative resultance. The subsection ends with an introduction to the concepts of extrinsic final value and includes some relevant remarks about the general kinds of objects that can possibly be good or bad for their own sakes.

Another view to be discussed is the fitting-attitude analysis, which states that final value is defined in terms of the presence of reasons for attitudes. This means that when an object is good or bad for its own sake, this is equivalent to there being reasons to hold positive or negative attitudes toward the object for its own sake. The second subsection illustrates the analysis by laying out what is supposed to be its main benefits and difficulties, including the famous wrong-kind of reasons problem. The problem involves doubt about the reduction of final value to the presence of reasons. Doubt crops up because there seem to be conceivable cases where we are given reasons to desire things that are not desirable, love things that are not lovable, abhor things that are not abhorrent, to hate things that are not hateful, and so on.

---

12 It would obviously be impossible to explicate and motivate all the philosophical assumptions that go into this work. What follows is merely a selection, based on what must be explicitly referenced and relied upon. This excludes some important assumptions that might nonetheless be deemed controversial by contemporary philosophers. Among them is the assumption of cognitivism, which states that statements expressing value judgments have an assertive illocutionary character. In other words, statements expressing value judgments aim at truth and falsehood in a strict sense, although they may fail to be either true or false (Joyce, 2001, pp. 8–9). This view is contrasted with that of non-cognitivism, according to which statements expressing value judgments have an expressive, commissive, or directive illocutionary character. On this view, statements expressing value judgments do not aim at truth and falsehood in a strict sense but are instead meant to express attitudes and commitments on the part of the speaker, or to influence the attitudes, choices, and behavior of other people (e.g., Barnes, 1933; Carnap, 1937; Stevenson, 1937; Ayer, 1952; Hare, 1952; Hägerström, 1953; Blackburn, 1984, 1993; Gibbard, 1990, 2003).
and so on. Two of the more influential responses to the problem will be presented, along with the reasons why they have been thought to fail. A seldom discussed solution to the wrong-kind of reasons problem will then be brought up, which suggests that the fitting-attitudes analysis could be made explicitly circular and have most of its benefits preserved. Some implications of this solution for our understanding of the attitudes that are involved in the fitting-attitudes analysis will also be highlighted. The subsection concludes with a discussion about the structure and nature of justificatory reasons.

Once they have been clarified and sufficiently elaborated upon, the combination of the various views discussed in the upcoming sections will, I think, allow us to see certain problems in a new light and offer up solutions to those problems that we may previously have been unaware of. This is not to say that the following will just consist in a straightforward application of the explanatory framework. Many specific arguments and theories will be discussed that are relatively unaffected by the combination of the views in question. Nevertheless, the explanatory framework could still be understood as constituting the glue of this work, or the red thread that ties it all together, since the views in question will be brought to bear on many of the key issues that are raised within it. For example, the framework will play a significant role in the attempts to defend the concept of final value, to account for the concept of background conditions, to explain the derivative way in which some objects can be good or bad for their own sakes, and to develop a satisfactory interpretation of the constitutive grounds of final value. We shall start, as is the custom, at the beginning, with the idea that normative properties generally and final value specifically must be taken to depend upon the presence of other factors.

**On Normative Explanations**

*Dependence* will here be understood as the broadest relation that involves everything that makes a difference for final value. The *dependence base* of final value is understood in a similar vein, as the combination of factors that figure in a complete explanation as to how an object ends up being good or bad for its own sake. In a complete explanation of this kind there may be factors belonging to several types. The reason we will, for the most part, ignore causal factors is that we will be discussing explanations of an exclusively synchronic character. When we ask a person why a painting is beautiful, they might give a causal story about how it originally came to be endowed with the...
involve, or subsume, several distinct relations. For practical purposes, we may assume that a complete explanation features factors belonging to roughly two distinct types: The first and arguably most central will be talked about in terms of evaluative resultance. This is the relation between final value and its makers, i.e. the specific set of factors we would first point to when asked to explain why an object has final value. The resultance base of final value is understood in a similar vein, as the specific combinations of factors in virtue of which, or as a result of which, an object ends up with final value. The second will be talked about in terms of background conditions and constitutive grounds. The latter relations must be distinguished later, but for now both can be taken to involve the factors responsible for other factors featuring in resultance bases.14

This way of seeing things entails that when we ask how objects end up being good or bad for their own sakes, we can have several different things in mind: On the one hand, we may be interested to hear the complete normative explanation, which considers any factor that might be relevant to the presence of final value. It is safe to say that such complete explanations are rarely represented in our evaluative talk and practices. The dependence of final value is an explanatory relation that is primarily invoked in philosophy and the more systematic parts of second-order value theory. On the other hand, we may also be interested in narrow explanations that only consider the factors in virtue of wonderful properties that it has. In many cases, however, such a story would be entirely beside the point. Instead, we might be interested to hear an explanation in terms of the properties that it has now, i.e. the features that underlie the beauty of the painting in a concurrent sense. Furthermore, causal explanations are associated with their own special problems in philosophy that we would do well to steer clear of, although there will unfortunately be a few instances where we have no choice but to briefly consider them.

---

14 Exactly in what way they are so responsible cannot be explained much more clearly at this stage, since this is a point of contention which the book aims to address in chapter three. Also, it might be thought that in addition to the two types of factors that have just been described, there is a third one which features prominently in full explanations for final value. I have in mind the factors that are responsible for the instantiation of the other factors. When a certain factor either makes an object have final value or is responsible for something else making an object have final value, then there might be yet another thing present which explains why the factor in question exists to begin with. For example, when an object is good for its own sake, then, in addition to the properties that explain why it is good for its own sake, there might also be factors that explain why the properties that make it good for its own sake are instantiated in the first place. These mediate instantiators, as it were, will not be discussed at any length in what follows. Instead, it will be assumed that factors playing this role involve a sufficiently similar relation to the one between final value and its makers that most of what is true of the latter will also be true of the former. Another reason why we do not pay as much attention to mediate instantiators is that they do not seem very relevant from the point of view of the similarity intuition, which, as we shall see, seems implicit in the dependence of final value.
which, or as a result of which, objects end up with final value. Such narrow explanations are by contrast frequently used in our everyday lives. Although narrow explanations may occasionally involve some background conditions, peripheral factors of this type tend to be treated as if they were unnecessary to mention. Here we will be stricter and simply assume that they are always irrelevant to narrow explanations.\textsuperscript{15} Now, distinguishing the dependence of final value from evaluative resultance invites another ambiguity in the concept of explanation which we would do well to consider before continuing.

Firstly, there is a practical and epistemic concept of explanation, understood by reference to the act of telling a story about why the world is the way it is. For example, this concept is used whenever people say of neuroscientists that they are trying to explain how certain brain processes give rise to conscious experience. Secondly, there is also a metaphysical concept of explanation, which is understood by reference to the circumstance of one thing being prior or being responsible for another.\textsuperscript{16} This concept is arguably used whenever cognitive neuroscientists say of certain brain processes that they give rise to conscious experience. Obviously, the relations we will be interested in here, including the dependence of final value and evaluative resultance, are assumed to be explanatory in the second sense. This means that when a final value is judged to either depend upon or result from the presence of certain factors, we commit ourselves to seeing the world as having a certain order and structure. In other words, the factors in question are judged to be prior or responsible for the instantiation of the relevant final value. To make this clearer, it might help to briefly consider some of the basic intuitions that tend to make us cognizant of the relations in question in the first place. Let us start out by looking at the broad relation involved in the dependence of final value upon other things.

Suppose we are walking through a museum when we overhear a friend say of two numerically distinct objects, such as a couple of paintings, that one of them is good for its own sake and that the other is bad for its own sake. This surprises us, let us say, because the paintings appear to us almost identical: They feature the same colors and shapes, are hung on the same walls, and are even subject to the same type of lighting. As far as we are aware, they even have similar historical features, being painted by the same painter at roughly the same time.

\textsuperscript{15} Some worries with this assumption are reconsidered later, when we begin to look at the deflationary view of background conditions.

\textsuperscript{16} Practical and epistemic notions of explanation are here treated as being equivalent just for the sake of making comparisons to the metaphysical notion of explanation, but some philosophers, like Broome (2013, p. 48), have emphasized that they are different.
Annoyingly, when we point this out to our friend, she just nods in agreement: The paintings are completely identical in every relevant respect, she says, save for the fact that one painting is good for its own sake and the other painting is bad for its own sake. This seems like a very strange answer, so we may imagine ourselves pressing the issue further and asking our friend to explain, in as broad terms as she can, why she thinks that one of the paintings is good for its own sake. Her answer is that nothing in particular makes it so. The one painting just happens to be endowed with a positive final value, while the other painting is not. The problem is that the value judgments expressed by our friend in this case run contrary to at least two deeply ingrained intuitions that many of us seem to share about final value and its dependence.

The first intuition states that whenever an object has final value, this cannot be a matter of accident. It is necessarily the case that if an object has final value, then the object must have other properties that explain, in a broad sense, why this is so. We are of course allowed to be unsure about what the relevant explanation is, but we know that there must be one. The second intuition is that two objects could not possibly differ only in respect of their final value. It is necessarily the case that if two objects differ in terms of their final value, then the objects must be different in some other way as well. As before, we may be unsure what the relevant difference between the two objects consists in, but we know that in principle there must be one. If this is on the right lines, then the question is invited whether the intuitions are of a first-order or second-order character. When our friend expresses value judgments contrary to the two intuitions, she seems not only to be wrong about some first-order issue but reveals herself to be in the grips of conceptual confusion. She is simply not a

17 Rønnow-Rasmussen makes the same point by saying that value “is not like butterflies that merely happen to settle on a flower” (2006, p. 104). This is part of the shared ground on which nearly all discussions in value take place. This is part of the reason why it has even become “a commonplace to observe that it is a commonplace to observe” that value supervenes upon other things (Majors, 2009, p. 29). See also Hare (1984, p. 1), Brink (1989, p. 160), Smith (1994, p. 20), Jackson (1998, p. 118), Shafer-Landau (2003, p. 77), and Strandberg (2004, p. 220). However, see also Sturgeon (2009) who argues that there is no universal understanding of dependence and so, on closer inspection, the shared ground on which we all stand turns out to have quite a few cracks in it.

18 We might find that the explanation itself has evaluative content, in that the value of an object depends upon the presence of another value. For example, perhaps a person being good in the sense of being admirable can come to depend on her being good in the sense of being respectable. We might also find that the explanation has purely nonnormative content, in that the value of an object can come to depend upon the nonnormative properties that the object has. For example, perhaps a person being good in the sense of being brave can come to depend on her being willing to face risks to attain the ends she perceives to be a worthy.
competent user of the concept of final value. This suggests that the intuitions in question are themselves conceptual and, consequently, that the view that final value depends upon other things is a second-order one. It is important to note that matters appear different when we narrow our focus and look at the same hypothetical case from the perspective of evaluative resultance. It is true, of course, that when an object has final value, there must also be something in virtue of which, or as a result of which, the object is good or bad for its own sake. The second intuition, however, does not seem to necessarily hold true when we consider only the makers of final value. Two objects could conceivably differ in terms of their final value even if they are completely identical in terms of the factors in virtue of which, or as a result of which, they have final value. From the perspective of evaluative resultance, then, the guarantee of similarity is either false or only true as a matter of substantive fact. It becomes a matter for first-order theory to find out whether it holds or not. The reason, to reiterate, is that there may be factors other than value-makers present in the complete explanation which affect the relevant final value. Background conditions and constitutive grounds are the paradigmatic examples, since they explain why other factors take part in the resultance base for final value to begin with, without thereby taking part in it themselves. If we interpret the value judgments expressed by our friend against the background of evaluative resultance, then, we can be far more forgiving.

We shall later be discussing the dependence of final value and the relation of evaluative resultance in further detail. For now, though, the most important thing to take away from all this is that the relations in question are invoked at distinct levels of explanation. The dependence of final value involves a very broad sort of explanation while evaluative resultance involves a very narrow one. The former is mainly of a philosophical and second-order interest, while the latter is the focus of our first-order everyday talk and practices. This is also reflected, to a degree, in the kinds of intuitions that we have looked at and that have motivated philosophers to carefully consider the relations in the first place. Now, complicating matters somewhat is the fact that there is a completely different sense in which normative explanations of this sort can be said to vary in width. This sense is best understood by reference to the distinction between extrinsic and intrinsic features. Before we move on, we should also consider the

---

19 The formulation of these intuitions invites several tough questions, among which is the question whether there is an analysis of dependence that can adequately capture them. This question is difficult enough that a detailed treatment of it will have to wait until the third chapter of this work.
meaning of this technical distinction and some of the controversies surrounding it, since it will prove very relevant to the understanding of final value that will be employed in the upcoming chapters.

On the Possibility of Extrinsic Final Value

The concept of final value has historically been tied to a great many notions, presumably under the assumption that they are meant to pick out the same thing. Most notably, it has been commonplace to suppose that if an object is good or bad for its own sake, then this must be explained by reference to its intrinsic features. These are the properties that an object has irrespective of the context of its occurrence, or as a result of its own inherent nature. Moore therefore said that the concept of intrinsic value is “that which we use when we say of an experience which we have had that it ‘was worth having for its own sake’” (1952, p. 94). So common was this supposition that it was not until recently that philosophers began speaking of final value and intrinsic value as if they were indeed different things. This more nuanced turn was motivated by two considerations, one of which revolves around substantive examples.

20 For more about the different value notions that have traditionally been talked about in conjunction with final value, see Feldman (1998). It should be admitted that the distinction between intrinsic and extrinsic features is a difficult one to understand. However, as a general earmark, it is helpful to think of extrinsic features as being those that an object has partly in virtue of the context in which the object is found. Intrinsic features, by contrast, are those that are had by an object irrespective of the circumstance of its occurrence. The distinction between intrinsic and extrinsic features has often been conflated with the distinction between essential and accidental features, respectively. This is a mistake. The redness of a rose is an intrinsic feature of it, but it is not an essential feature. If it is exposed to too much sunlight, it can become withered and brown. This means that when final value is intrinsic, it need not be essential either. As the flower withers and becomes brown, it can lose some of its final value.

21 This is why philosophers like Moore (1993/1903, pp. 142–147) thought that we can find out whether an object has final value by imagining it in isolation from the context of its occurrence. For more about the isolation test, its variations, and its weaknesses, see, e.g., Broad (1961), Chisholm (1978, 1981), Bodanszky & Conee (1981), and Lemos (1994, pp. 3–19).

22 It is uncertain whether this means that the distinction between, say, final and instrumental value should be treated as entirely orthogonal to the distinction between intrinsic and extrinsic value. For although it may seem reasonable to assume that final value can come in both the intrinsic and the extrinsic variety, the same does not seem to be true for instrumental value. In particular, it seems to me difficult to conceive of a convincing example where an object is good or bad for the sake of its effects, but entirely as a result of the intrinsic features possessed by the object. Dorsey (2012) offers some examples along these lines which are criticized in Rønnow-Rasmussen (2015). Another interesting observation in the vicinity of this topic, originally made by Kagan (1998, pp. 283–284), is that it might be possible for an object to be good or
The Coin: Collectors might judge that a specific coin is good for its own sake. However, they might also think that its intrinsic features do not make a significant contribution toward its final value. Neither its design nor its materials seem to make it any better for its own sake than many other worthless coins that can be found in the world. What makes this specific coin good for its own sake is rather something to do with its extrinsic features, such as its property of being one of the rarest coins in existence.24

The Pen: When George Washington signed the United States Constitution he used a pen that many historians would value. The pen might be ascribed instrumental value, since it can be used to form emotional connections to past events. However, the pen might also be ascribed with final value, precisely because of its historical significance. The pen might be good for its own sake in virtue of its extrinsic feature of having been used by George Washington to sign the United States Constitution.25

The Wilderness: People often value environmental sites, such as forests and rivers. They might judge such sites to be instrumentally good, since they can offer experiences of relaxation and natural beauty. However, it seems possible for a forest to be good for its own sake, at least partly in virtue of the limited connection it has to human activities to begin with. If this is right, then the wilderness can owe its final value to its extrinsic feature of being relatively untouched by human hands.26
It is important to keep in mind that in second-order value theory substantive examples like these are often meant to be examples of examples. This means that we do not have to take the value judgments involved in any one of these cases at face value. It might be that neither rare coins, historically significant pens, nor the untouched wilderness end up being good or bad for their own sakes. What is important is that none of the value judgments invited by the examples seem conceptually confused. A person might be coherent even if they did think rare coins, historically significant pens, and the untouched wilderness end up being good or bad for their own sakes. This is enough to indicate that the concept of final value is distinct from the concept of intrinsic value, with which it is often conflated. If they were not distinct, then we would not merely be positioned to think that the substantive examples get things wrong from a first-order perspective. We would even have to conclude that, strictly speaking, the substantive examples do not make much sense.

It is also important to keep in mind that by separating the dependence of final value from evaluative resultant, an ambiguity is introduced into the present discussion. An object can have extrinsic final value in the sense that a complete explanation of it will refer to extrinsic features, or in the stronger sense that a narrow explanation of it will refer to extrinsic features. This is an ambiguity that we should be able to live with, but it is also one which needs to be kept firmly in mind. Whenever the possibility of extrinsic or intrinsic final value is addressed, we need to be entirely clear about which sense is relevant to the discussion at hand. For example, if we see the extrinsic factors mentioned in the cases above as constituting background conditions for the final value of the relevant objects, then we have in mind the wide sense of extrinsicality. If we see the extrinsic factors in question as constituting makers of final value, then we have in mind a narrow sense of extrinsicality instead.

The second consideration which has motivated the more nuanced turn regarding explanations for final value is not so much motivated by an appeal to substantive examples. Instead, it is motivated by the more general suggestion that there is something important which separates questions about why an object has the value that it has and in what way the object has value (cf. Korsgaard, 1983, p. 78). The former questions are concerned with what explains the presence of the relevant value, while the latter questions have more to do with the type of value that ends up being present. Now, some philosophers remain unconvinced that these kinds of questions are practically distinct. They tend to couch their objections in terms of the kinds of things that could possibly end up being the bearers of final value. We have so far left it entirely open that all manner of things could be good or bad for their own sakes. However, there is
a pervasive tendency within the field to ascribe final value only to certain categories of objects, such as facts and states of affairs (e.g. Ross, 1930; Harman, 1960; Chisholm, 1968; Quinn, 1974; Oldfield, 1977; Quinn, 1977; Carlson, 1995; Feldman, 1997; Bradley, 1998; Zimmerman, 2001).

Roughly, states of affairs are entities with a propositional character that can be picked out by that-clauses. For example, it is a state of affairs that the cat sits on the mat. We may say that when a state of affairs like this obtains and it is true that the cat sits on the mat, it also becomes a fact. It will be a fact, in other words, that the cat sits on the mat. The view now under consideration, then, states that only entities of a propositional character, like the fact that the cat sits on the mat, could be good or bad for their own sake. By itself this does not show that all final value must be intrinsic, since states of affairs and facts can also have extrinsic features. Nevertheless, rejecting pluralism in this way invites an obvious argument going in this general direction: When the coin is judged to be good for its own sake in virtue of being rare, it might be claimed that what has final value is the fact that the coin is rare. When the pen is judged to be good for its own sake in virtue of having been used by Washington, it might be claimed that what has final value is the fact that the pen was used by Washington. Finally, when the wilderness is judged to be good for its own sake in virtue of being untouched by human hands, it might be claimed that what has final value is the fact that the wilderness is untouched by human hands.

The view that final value must be intrinsic, then, is given most of its theoretical support from a rejection of pluralism about the bearers of final value.27 Unfortunately, it turns out that the grounds for this rejection are rather shaky. The maneuver described at the end of the last paragraph has an undeniable air of revisionism to it. The question is why the maneuver ought to be employed at all: Why not just leave everything as is and assume that all manner of things could be the bearers of final value? It might be true that if all we had to deal with were facts or states of affairs, formal work on final value

---

27 Anderson (1993, p. 20) rightly points out that the kinds of attitudes that are rational to direct toward things that are valuable are not always propositional. That is to say, they do not always have facts or states of affairs as their intentional objects, but are instead directed toward something concrete, like paintings or human beings. See also Swanton (1995) and Baron (1997) for observations of this sort. Rabinowicz & Rønnow-Rasmussen (2003, pp. 391–392) agree and so speculate that another reason why some philosophers might be tempted to only ascribe final value to facts or states of affairs is that they, for some reason or other, want to restrict the scope of attitudes that are relevant to value to things like preferences, wishes, and hopes. Again, it is not easy to see the rationale for such a restriction, aside from the fact that it might make things easier from a methodological standpoint.
might be made easier. However, methodological benefits like these say less about how things are and more about how we would like them to be. Since there are coherent examples of concrete objects that have final value, we need a reason to reject this possibility aside from the fact that such a rejection would make the lives of philosophers less difficult. Michael J. Zimmerman (2001) is among the very few who have attempted to properly meet this challenge. His arguments are sophisticated and interesting in their details, but many of them seem to me to boil down to a suggestion centering on the concept of derivative value.

The suggestion is that whenever concrete objects appear to be good or bad for their own sakes, they really have a derived form of value. This is so, Zimmerman says, because the value of concrete objects is always to be explained in terms of the value of the facts that they are included in. He (2001, pp. 36–38, 45–46) formulates several substantive examples that are meant to move our intuitions in the direction of his suggestion, but we will not concern ourselves with those now. The most serious problem with the suggestion does not have to do with any of the substantive examples that he offers in its support, although they too have been criticized and shown to be subject to some forceful objections (e.g. Rabinowicz & Rønnow-Rasmussen, 2003, pp. 393–395). Instead, the most severe problem is that the suggestion comes very close to begging the question against proponents of pluralism. Those who are tempted to deny that final value is always intrinsic are just as likely to deny that final value cannot be derivative. 28

The concept of derivative value is a difficult one to which we shall have to return. For now, though, we may suppose that an object has derivative value if, and only if, it owes its value to the presence of another normative property. This is enough indication that the concept of derivative value is more relevant to questions about why an object has value, rather than in what way the object has value (Rabinowicz & Rønnow-Rasmussen, 2003, pp. 392–393).

We may first respond to Zimmerman, then, by saying that it is doubtful that he is right about the value of concrete objects always being derived from the final value of facts or states of affairs. In some cases, the opposite might be true. Regardless, it does not follow that concrete objects cannot therefore have final value. Just as there can be extrinsic final value, there can be derivative final value. In fact, derivative value will often be a sort of extrinsic value. The value of an object is extrinsic if, and only if, the object owes the value at least partly to the context of its occurrence. The value of the object could also be said to be derivative if, and only if, said context of occurrence includes among its relevant

---

28 For a proponent of extrinsic final value who denies the possibility of derivative final value, see Olson (2004, pp. 34–35).
parts the presence of another normative property. 29 We shall not discuss these issues further here. It will be assumed throughout the following work that the concept of final value is different from the concept of intrinsic value. It will also be assumed that final value can accrue to all manner of objects, be they concrete or abstract. This assumption might turn out to be wrong, but this, in my view, remains to be shown.

On the Fitting-Attitudes Analysis

We shall here be assuming that the evaluative landscape, of which the concept of final value forms an integral part, is not an independent land, completely divorced from the wider normative landscape. The opposite is in fact true. It seems reasonable to suggest that final value itself, at least as it tends to figure in our evaluative talk and practices, is normative. Not surprisingly, there is significant controversy among contemporary philosophers as to how this seemingly simple observation should be understood. The theory relied upon in this work is a rather radical one, saying that value generally and final value particularly should be analyzed directly in terms of the normative. For something to be good or bad, then, is, at least in part, for it to have normative implications of a certain sort. The fitting-attitudes analysis of value cashes this out in a more precise manner by stating that for an object to have value is for it to be the fitting object of some type of attitude. When applied to the concept of final value, what results from this pattern of analysis are the following:

\[ F1: \quad x \text{ is finally good } =_{df} \text{ it is fitting to favor } x \text{ for its own sake.} \]

\[ F2: \quad x \text{ is finally bad } =_{df} \text{ it is fitting to disfavor } x \text{ for its own sake.} \]

Favoring should hereafter be understood as a placeholder for a wide range of different positive responses, the character of which will vary depending on the specific type of value that is being analyzed: iff the object is respectable, then it will be fitting to respect it; iff the object is admirable, then it will be fitting to

---

29 That is not to say that derivative value is always extrinsic. Derivative value is not extrinsic value in cases where it is derived from the presence of another value that is also intrinsic to the relevant object. The value of a painting may be instantiated in virtue of, or as a result of, the value that accrues to the signature on the painting. The value of a complex fact may be instantiated in virtue of, or as a result of, the value that accrues to one of its constituent facts.
admire it; iff the object is lovable, then it will be fitting to love it; and so on. When the fitting favoring does not seem to be constituted by a motivational attitude of the sort just mentioned, what is most relevant is presumably some kind of behavioral disposition, like the disposition to protect, promote, or preserve the object. Disfavoring should of course also be understood as a placeholder in this sense, but one which stands for a wide range of different negative responses: iff the object is condemnable, then it will be fitting to condemn it; iff the object is deplorable, then it will be fitting to deplore it; iff the object is detestable, then it will be fitting to detest it; and so on. Once again, when the fitting disfavoring does not seem to be constituted by a motivational attitude, what is most relevant is presumably some kind of behavioral disposition, like the disposition to prevent, suppress, or destroy the object under consideration. The same pattern of analysis can of course be applied to comparative cases as well:

$$F3: \quad x \text{ is finally better than } y \quad (y \text{ is finally worse than } x) =_{df} \text{ it is fitting to prefer } x \text{ over } y \text{ for its own sake.} \quad 31$$

$$F4: \quad x \text{ is finally equal to } y \quad (y \text{ is finally equal to } x) =_{df} \text{ it is fitting to be indifferent between } x \text{ and } y. \quad 32$$

The fitting-attitudes analysis represents a tradition with a surprisingly long history in value theory. One of the earliest proponents was Franz Brentano in the nineteenth century who suggested that, “the good is that which is worthy of

---

30 I will for the most part remain non-committal on the matter of whether attitudes can in the end be analyzed in terms of dispositions, though I admit to skepticism about this possibility.

31 John Broome (2004, pp. 50–51) has suggested that for x to be better than y just is for x to have more value than y. This understanding seems especially fitting when we consider other comparatives: For x to be taller than y just is for x to have more height than y; for x to be heavier than y just is for x to have more weight than y; for x to be older than y just is for x to have a greater age than y; and so on. While it may appear very reasonable to analyze the notion of x being better than y in terms of fitting preferences, the notion of x having more value than y recommends a completely different approach. This approach eschews reliance upon dyadic attitudes like preferences and instead understands the comparative case in terms of the fittingness of monadic attitudes. The idea is adopted by Rabinowicz (2012) who argues that when x is better than y, then it is fitting to have stronger attitudes toward x than y. Perhaps this monadic approach to evaluative comparatives must be adopted if we are to be able to capture what it would mean for an object x to be better than y to a certain degree. For example, when we rely on preferences we seem to lose the ability to straightforwardly account for what it would mean for x to be twice as good as y.

32 See also Green Werkmäster & Andersson (unpublished manuscript).
love, that which can be loved with a love that is correct” (1969/1889, p. 18). Since then it has been defended in some form by a large group of philosophers which includes C. D. Broad (1930), A. C. Ewing (1947), Richard Brandt (1946), John McDowell (1985), Roderick Chisholm (1986), W. D. Falk (1986), David Wiggins (1987), Allan Gibbard (1990), Elisabeth Anderson (1993), Noah Lemos (1994), Thomas Scanlon (1998), John Skorupski (2010), and many others. Although their accounts tend to differ in their details, what unites them is that they analyze value in terms of a response, on the one hand, and some deontic notion, on the other hand. The reason for the longevity and continuing popularity of the tradition, aside from the immediate intuitive appeal that it seems to hold for many philosophers, is that it brings with it several theoretical advantages. It is worth pausing for a moment and taking stock of some of those advantages, that we may clearly establish what is at stake.

First, the fitting-attitudes analysis promises significant conceptual gain. Second-order value theory concerns itself with countless different concepts, only a handful of which have so far been mentioned. The fitting-attitudes analysis promises to bring some order to the apparent chaos by suggesting that at least some of those concepts, namely the evaluative ones, can be partly understood in terms of certain other concepts, namely reasons (more specifically, reasons for attitudes). In so doing, it reduces the overall complexity that seems to characterize the normative domain, thereby relieving it of some of its apparent mystery. Second, the fitting-attitudes analysis ties value to the presence of attitudes in a way that seems relatively plausible. Just as value is thought to be inherently normative, it seems obvious to many of us that it has some connection to attitudes. The fitting-attitudes analysis gives us a picture of what

33 For more on these sources and their different understandings of the fitting-attitudes analysis, see Rabinowicz & Rønnow-Rasmussen (2004, pp. 394–400, with accompanying footnotes).
35 When teaching students about the merits of the fitting-attitudes analysis, it is not uncommon to state that the analysis demystifies final value. Given the mysteries associated with reasons themselves, it may seem doubtful whether this is true. However, the fitting-attitudes analysis does make the apparent normativity of value less mysterious, even if normativity itself remains strange (e.g. Rabinowicz & Rønnow-Rasmussen, 2004, p. 400). What is more, it seems right to say that by wedding value to reasons, the mystery that is associated with the normative domain as such is mitigated. Whenever a structure is imposed on a set of elements, then it seems reasonable to say that the set is illuminated and made less mysterious, even if the elements themselves retain some of their mystery. This is enough to make the fitting-attitudes analysis seem attractive to a philosopher.
this connection looks like. It is not necessarily the case that value entails the presence of attitudes. What is guaranteed by value, conceptually speaking, is only that that an attitude would be fitting. In this way, attitudes are placed at the center of our understanding of what final value is.

Third, the fitting-attitudes analysis appears to be sufficiently neutral in that it does not force us to choose sides in regard to too many other issues in philosophy. Analyses of value tend to run the risk 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 relations (Rabinowicz & Rønnow-Rasmussen, 2004, p. 400). Third and finally, the fitting-attitudes analysis promises significant epistemic gain. It is informative in that it helps to shed light on some old and familiar notions. Indeed, the following work could be seen as offering a demonstration of this capacity, since it applies the fitting-attitudes analysis in order to accomplish two goals: the first is to illuminate concepts like final value, instrumental value, kind-value, and personal value; the second goal is to illuminate the dependencies and relations of the former by looking at the phenomenon of background conditions, constitutive grounds, and so on.

Two Problems for the Fitting-Attitudes Analysis

For all its advantages, the fitting-attitudes analysis of value is unfortunately faced with some very difficult objections. I shall briefly mention two of the most difficult ones and explain how I think that they should be dealt with. The first objection has to do with the attitudes that are appealed to by the fitting-attitudes analysis (Rabinowicz & Rønnow-Rasmussen, 2004, p 401): What does it mean to favor an object as opposed to disfavoring that same object? How do the attitudes that are called for by objects with final value differ from the attitudes that are called for by objects with other types of value? Without answers to these

---

36 Some philosophers worry that the fitting-attitudes analysis settles certain first-order debates about the soundness of theories like consequentialism (Dancy, 2000). However, those worries have since been put to rest (Olson, 2007).

37 Although, see Reisner (2015), who argues that the fitting-attitudes analysis is not informative.

38 The fitting-attitudes analysis has been used to develop formal models that capture more elusive comparatives, including the elusive parity relation (e.g. Gert, 2004; Rabinowicz, 2008, 2012).
questions, the fitting-attitudes analysis might not seem so illuminating after all. The first thing to say here is that the conceptual project of attaining a better understanding of attitudes is a general one. Many accounts both in and outside of philosophy appeal to attitudes of this kind and so in so far as we lack a firm grasp of their nature, this is a respect in which it is fair to say that everybody hurts. Furthermore, it seems to me that the fitting-attitudes analysis could be illuminating even without answers to these types of questions. To know that value is tied to reasons and attitudes in the indicated way tightens our grip on it, even if attitudes themselves remain very mysterious.

We shall soon reconsider the objection just mentioned, but we will first turn to another objection that has received far more constructive attention within the literature. The objection states that the fitting-attitudes analysis fails since there are circumstances where the entailment from reasons for attitudes to value does not hold. There can be reasons to love that which is not lovable, to respect that which is not respectable, to desire that which is not desirable, to condemn that which is not condemnable, to deplore that which is not deplorable, and to detest that which is not detestable. More generally put, there can be reasons to favor that which is not good, just as there can be reasons to disfavor that which is not bad. The previous problem had to do with the proper understanding of attitudes, but the one just described instead focuses on the normative component invoked by the fitting-attitudes analysis. Within the first-order domain, this issue has quite a bit of history to it. In the nineteenth century, Henry Sidgwick (1981/1907, p. 48) were among the philosophers to observe that even if it turns out that happiness is the only thing that has final value, we can nonetheless have reasons to favor other things for their own sakes. The explanation for this is that if we do not favor other things for their own sakes, it becomes more unlikely that happiness is eventually attained. In the context of second-order value theory as it is here understood, the problem is often illustrated with examples involving violence or the threat of violence. The following versions of the examples are among the most famous:

*The Mud Saucer:* An Evil Demon threatens us, on pain of torture and death, to admire a saucer of mud for its own sake. In spite of the saucer of mud being bad or having no value at all, it seems to make sense that the Demon’s threats give us reasons to admire the saucer of mud for its own sake. If this is right, then there can be reasons for attitudes that apparently do not entail the presence of value.
The Evil Demon: An Evil Demon threatens us, on pain of torture and death, to admire him for his own sake, on account of his threatening ways. Despite the Evil Demon being entirely bad or having no value at all, it makes sense to say that his threats give us reasons to admire him for his own sake. If this is right, then there are reasons for attitudes that apparently do not entail the presence of value.39

There have been several proposed solutions to the problem (e.g. Olson, 2004; Stratton-Lake, 2005; Danielsson & Olson, 2007; Persson, 2007; Lang, 2008; Schroeder, 2010; Way, 2012). We shall not go through all these solutions but will look very briefly at some of the more popular ones to come out of the literature and note their differences and similarities. One solution is of course to deny that there is a problem in the first place. The most influential suggestion to this effect, inspired by the writings of Derek Parfit (2001) and most enthusiastically championed by John Skorupski (2010), relies upon a distinction between epistemic, evaluative, and practical reasons. In short, epistemic reasons are reasons for beliefs, evaluative reasons are reasons for attitudes, and practical reasons are reasons for action. The idea is that the reasons involved in the kinds of cases we have been looking at are practical rather than evaluative. This means that they cannot be reasons for attitudes in the first place. At most, they are reasons for actions, such as adopting an attitude or making it the case that we have an attitude. That the relevant reasons are practical is supposedly indicated by their source. When the Evil Demon forces us to favor either the saucer of mud or him, we have reasons to comply only in so far as having the relevant attitudes will help us avoid his punishment.40

39 Examples involving an Evil Demon were first formulated by Crisp (2000, p. 459). Similar examples can be found in the early discussions of D’arms & Jacobson (2000), although they were more concerned with reasons against favoring objects that are nonetheless good or bad. These variations on the Evil Demon example come from the discussions of Rabinowicz & Ronnow-Rasmussen (2004, p. 407). The examples are useful because of their clarity and force, but they are of course slightly unrealistic, to put things mildly. Fortunately, versions of the examples that are more down to earth are not difficult to formulate. A parent can have reasons to love her child more than any other even though the child is not more lovable (Orsi, 2015, pp. 144–145). A person can have reasons not to be amused by a joke even though the joke is amusing (Gaut, 2007, p. 241; Orsi, 2015, pp. 144–145). A person can have reasons to admire the suggestion of an employer even though the suggestion is not very admirable. The problem with these sorts of examples, as Wlodek Rabinowicz has pointed out (in private communication), is that it might be argued that the fitting-attitudes analysis only concerns itself with the reasons there are for everyone.

40 Parfit (2001) develops this further and weds his view of practical reasons to a substantive story about their source. He refers to reasons that are facts about attitudes as ‘state-given’ and
The main objection against this solution to the problem is that we have no way of telling exactly when a reason is evaluative and when it is practical. Obviously, we cannot, on pain of begging the question, simply assume that all cases that seem to tell against the fitting-attitudes analysis must be cases where we are given practical reasons rather than evaluative reasons. Without a systematic way of distinguishing the two, there is also no way to gain leverage against philosophers who continue to insist that the cases we have looked at do involve evaluative reasons and not just practical reasons. Rabinowicz & Rønnow-Rasmussen (2004, pp. 412–414) are among those philosophers. They claim that when the Evil Demon threatens us to favor him for his own sake, it seems intuitive that there should be an evaluative reason to favor him for his own sake. The Evil Demon is not impressed by mere attempts to favor him, after all, but only cares if we successfully manage to do so. In so far as we have a reason to adopt the attitude, then, this might be because there are reasons to hold the attitude. Here, there is a significant risk that the discussion will turn into mere table-thumping over substantive intuition.

reasons that are facts about objects as ‘object-given’. A reason to favor another attitude is state-given if it is constituted by facts about the favoring there are reasons to have and it is object-given if it is constituted by facts about the object of the favoring that there are reasons to have. Rabinowicz & Rønnow-Rasmussen (2004, pp. 406–407) have shown that the bigger problem is a different one, namely that for every fact about the properties of attitudes there is a corresponding fact about the properties of the intentional objects of the attitudes. For example, if a positive attitude directed at the Evil Demon has the property of being useful in the sense that it helps us avoid torture and death, then the Evil Demon has the property of being such that a positive attitude directed at him has the property of being useful in the sense that it helps us avoid torture and death. This leaves the door open for the view that where there are state-given reasons there are necessarily, or at least typically, object-given reasons as well. Another very big problem is that there do seem to be object-given reasons that are nonetheless reasons of the wrong kind. The reasons to love our own children more than those of others are potentially a case in point (Orsi, 2015, pp. 144–145). These reasons do not necessarily consist in facts about attitudes, but of facts about the parents, child, and their special relations.

41 This is reminiscent of debates in epistemology about Pascal’s wager (Orsi, 2015, p. 147). Pascal thought that we can have reasons to believe in the existence of God because of the risks involved in not doing so and being wrong. Here it seems perfectly rational to argue that the threat of being wrong does not give us reasons to believe in God. At most it gives us practical reasons to make ourselves believe in God, or something to that effect. Unfortunately, the analogy here has weaknesses. We know that the threat of being wrong is not a reason to believe in God, because it is not an evidential reason. In other words, that there is a significant risk in being wrong about the existence of God does not make it more likely that God exists. In the case of value, no similar line of reasoning is thought to be immediately invited, unless we make the analysis circular (Orsi, 2015, pp. 148–150). We could say that reasons for attitudes that are of the right kind are those that in some sense speak to the value of the relevant object. However, once we go down this route, there is, as we shall soon see, a much simpler way of avoiding the wrong-kind of reasons problem.
Another impulse might be to say that the value of objects can only correspond to reasons for attitudes that are facts about the intrinsic properties of said objects. Unfortunately, this solves the problem for intrinsic final value, but not for extrinsic final value (Rabinowicz & Rønnow-Rasmussen, 2004, pp. 408–409). Nevertheless, perhaps this line of response gets something right in that it at least acknowledges that the kinds of cases we have looked at do involve genuine evaluative reasons for attitudes. In this way, the approach tries to get around the problem, not by denying that it crops up in the first place, but by trying to distinguish reasons of the wrong kind from reasons of the right kind. The idea, then, is that value can be analyzed in terms of the presence of reasons for attitudes, but not all reasons for attitudes that happen to be present are relevant from an evaluative point of view. One of the more interesting suggestions to this effect, originally brought up and explored by Rabinowicz & Rønnow-Rasmussen (2004, pp. 414–418), proceeds by looking at the way things should figure in the intentional contents of fitting attitudes.

Their idea is that for a reason to be of the right kind, the fact that constitutes the reason needs to do two different jobs, one of which is justificatory and the other is representative.42 Suppose we favor an object for its own sake on account of certain properties that it has. If the fact that the object has these properties constitutes a reason to favor it, on the one hand, and if the fact that the object has these properties figures in the attitude as a reason to favor it, on the other hand, then the reason is of the right kind.43 This might help us get away from examples along the lines of the first one mentioned earlier, involving the saucer of mud. When we favor the saucer of mud in this example, we are being compelled to do so for the properties that the saucer of mud has. However, the

---

42 See also Cook (2008) for a defense and Rønnow-Rasmussen (2013) for some further elaborations upon the dual-role solution to the wrong-kind of reasons problem.

43 The cake can be cut in different ways here. Consider the fact that a specific concrete object has a certain property. For this fact to constitute a reason of the right kind two things are required of it, on the dual-role approach just explicated. The first is that the reason it constitutes must be a reason on the part of some agent to favor an object. The object that is favored need not be the same concrete object that is involved in the fact which constitutes the reason (while they are not the same in the Mud Saucer example, they are the same in the Evil Demon example). So far, then, everything looks much the same as it did above. It turns out, however, that the second requirement can be cashed out in several ways. Instead of saying that the fact must figure in the intentional content of the favoring as the thing that justifies the favoring, we can instead refer to something that is involved within the fact, as it were. For example, we might say that the property which is ascribed to the concrete object within the fact needs to figure in the intentional content of the relevant favoring as a value-maker. When you look at what they explicitly say, this seems more in line with the original vision of Rabinowicz & Rønnow-Rasmussen (2004, pp. 415–415).
fact that plays the justificatory role in this case, constituting the reason to favor the useless saucer of mud, has nothing to do with the saucer of mud. The reason is instead constituted by the fact that we have been threatened and we want to avoid punishment. We may therefore claim that the reason in question is not of the right kind to be relevant to the analysis. Unfortunately, even if the approach manages to be effective in this case, it is doubtful whether it really generalizes.

However, when we are compelled to favor the Evil Demon for his own sake, we are meant to do so because of the properties that the Evil Demon has. We are to admire him, let us say, because he happens to be so threatening to us (Rabinowicz & Rønnow-Rasmussen, 2004, pp. 419–420). In this case, the fact that the Evil Demon is threatening seems to do all the jobs that are required for it to be a reason of the right kind. The fact in question plays the relevant justificatory role and it is represented in the intentional content of the favoring as the very reason that justifies it. Attempts to modify the approach so that it can accommodate the second example have been made, as have endeavors to find alternative ways of distinguishing reasons of the right kind from reasons of the wrong kind (e.g. Olson, 2004; Stratton-Lake, 2005; Danielsson & Olson, 2007; Lang, 2008; Schroeder, 2010; Way, 2012). One approach that has not received much attention in the literature and that I would like to consider in the next subsection states that reasons for attitudes are of the right kind only if they figure in the intentional contents of said attitudes as value-makers. Furthermore, the reasons in question are to be grounded in facts about what actually makes the relevant objects good or bad for their own sakes. Relying upon these suggestions would make the fitting-attitudes analysis circular, but on my view the circularity in question is benign and ought to be welcomed.

Conceptual Circularity and the Analysis of Value

Much of the following work involves an exercise in conceptual analysis, a method which is markedly controversial in its characterization and application. On a very narrow understanding, conceptual analysis is the result of a method whereby one set of concepts is reduced to another set of concepts by formulating definitions in terms of necessary and sufficient conditions. My understanding of conceptual analysis tends to be more liberal and takes it to be, roughly, the activity of “trying to gain knowledge—philosophical or otherwise—which is based on conceptual competence” (Kipper, 2012, p. 9). The activity of conceptual analysis does not preclude the use of definitions in terms of necessary and sufficient conditions, but it treats these as means to an end. Definitions in
terms of necessary and sufficient conditions would be used because there is a desire to be as precise as possible, while we endeavor to elucidate the concepts represented in our everyday talk and practices. This is of course compatible with acknowledging that there are limits to how much definitions in terms of necessary and sufficient conditions can capture, and it does not commit us to the project of finding reductions.

Considering what has just been said, we might feel the temptation to simply acquiesce to the counterexamples mentioned earlier and claim that they represent the limits of the definitions afforded by the fitting-attitudes analysis. Here, no further adjustment can be made, no additional necessary condition added, that would help us get away from the cases represented by the counterexamples. There is only the choice of giving up on the definitions altogether or sacrificing the intuitions that stand in their way, thereby revising the very concepts we were hoping to better understand. The main trouble here is that it is very difficult to know exactly when such a limit has really been reached. There is a chance that continued attempts to adjust the fitting-attitudes analysis, even if they do not end up being entirely foolproof, can themselves teach us important lessons about the concept of final value and help us situate it more firmly within our broader conceptual schemes.44 Instead, the strategy that I

---

44 Analogies can perhaps be made here with discussions in other areas of philosophy, including epistemology. By formulating his famous objections to the analysis of knowledge in terms of justified and true belief, Gettier (1963) inspired a flurry of attempts to adjust the analysis, either by adding necessary conditions or by removing them (Ichikawa & Steup, 2017). Although there have been some promising candidates, no one attempt along these lines has enjoyed wide acceptance among philosophers within the field. However, that is not to suggest that the attempts have been completely unfruitful, for that would clearly be wrong. The debates that have taken place in the past fifty years about the correct analysis of knowledge have significantly deepened our understanding of concepts like knowledge, truth, justification, and belief. Furthermore, it is worth observing that in the process of looking for better conceptual analyses like these, we may also find necessary conditions that are at least better than those we had before. In this way, we can narrow down the types of problem cases that a given analysis happens to be burdened with, even though we cannot eliminate all such problem cases. Sandin makes the same point in defense of conceptual analysis when he writes: “Even if we were to accept that no such definition [in terms of necessary and sufficient conditions] is to be found, the activity of searching for such definitions need not be pointless. It might well be that we gain something else from the search. Here is one obvious example: We gain definitions that are better than the one we had before. Again, an analogy from science: It might be reasonable to go on searching for a true, complete description of a phenomenon, while admitting that there is no such thing. The descriptions obtained might still be better than the ones we decide to scrap” (2006, p. 31). I think that arguments along these lines can be developed further and constitute strong defenses of the method of conceptual analysis. It is a shame that they have not received as much attention in the literature as they deserve.
would advocate for involves giving up on the reductive aspects of the fitting-attitudes analysis and invite conceptual circularity.

Several philosophers have made suggestions along these lines, among them being Wiggins (1987), McDowell (1998), and Orsi (2015). The latter at one point considers a proposal according to which an object is finally good if, and only if, there are reasons to judge it to be good for its own sake. Conversely, an object is finally bad if, and only if, there are reasons to judge it to be bad for its own sake. According to Orsi (ibid, p. 157), the main problem with this account is that by formulating the fitting-attitudes analysis in terms of judgment, the reasons that are relevant to the understanding of value will all be of the epistemic sort. A related worry is that we may want different sorts of value to call for different sorts of responses. From this perspective, Wiggins (1987, p. 195) offers a more promising account, since he instead takes a wider view and states that the attitudes that are relevant to the case of value must be understood in partly evaluative terms. This is especially promising given the first objection mentioned in the previous section. The objection in question consists in two distinct challenges. Let us remind ourselves of what those challenges are.

One challenge is to explain how favoring an object differs from disfavoring it and another challenge is to explain how the attitudes that are called for by objects with final value are meant to differ from the attitudes that are called for by objects with other types of value. If we are not concerned with reducing value concepts, then several suggestions that take inspiration from Wiggins invite themselves. One natural suggestion along these lines would be that an object is favored only if it is represented in the intentional content of the attitude as being good. An object is disfavored only if it is represented in the intentional content of the attitude as being bad. Another suggestion along these lines would be that an object is favored for its own sake only if it is represented in the intentional content of the attitude as being finally good. An object is disfavored for its own sake only if it is represented in the intentional content of the attitude as being finally bad. Indeed, all of this makes it seem as if any philosopher trying to capture the distinct types of value that there are by relying on a corresponding set of properly discerning attitudes will be tempted toward an explicitly circular version of the fitting-attitudes analysis.

Unfortunately, although the circular fitting-attitudes analysis inspired by Wiggins comes some way in answering the first objection we considered, it does not suffice to save the fitting-attitudes analysis from the wrong-kind of reasons

---

However, I shall have to leave them aside for now. Some of these issues will be brought up and discussed in the appendix, when I discuss the issue of conceptual circularity and its virtues.
problem. From this perspective, both the account mentioned by Orsi and the one inspired by Wiggins seem doomed to fail. The reason is that their approach hinges upon the introduction of conceptual circularity within the intentional contents of the attitudes or judgments appealed to by the analysis. This leaves the door wide open for the wrong-kind of reasons problem. After all, the Evil Demon could simply state that he will punish us with severe torture and death unless we adopt the attitudes in question, wherein factors are represented in the way that we happen to think is required. For conceptual circularity to be helpful from the perspective of the wrong-kind of reasons problem, then, it needs to be introduced in a different manner that speaks more to the actual value that is being analyzed. One such strategy is to formulate the fitting-attitudes analysis in terms of the reasons that are grounded in the factors that also make the objects good or bad for their own sakes.45

At this stage we could take inspiration from the dual-role solution proposed by Rabinowicz & Rønnow-Rasmussen and combine the two kinds of conceptual circularity just described, thus incorporating the two distinct substantive requirements that have just been alluded to within our formulation of the fitting-attitudes analysis: One requirement would concern the normative role that the factors that ground the fittingness of attitudes must play, while the other requirement would concern the way that factors are to feature within the intentional contents of said attitudes. The result of this circular version of the dual-role approach would look something like this:

\[ F5: \quad \text{x is finally good} =_{df} \text{it is fitting to direct an attitude toward x such that i) the fittingness of the attitude is grounded in factors that make x good for its own sake, and ii) the attitude represents the factors that ground its fittingness as the things that make x good for its own sake.} \]

\[ F6: \quad \text{x is finally bad} =_{df} \text{it is fitting to direct an attitude toward x such that i) the fittingness of the attitude is grounded in factors that make x bad for its own sake, and ii) the attitude represents the factors that ground its fittingness as the things that make x bad for its own sake.} \]

45 This is not negated by the fact that the latter concept cannot itself be understood without reference to a value concept (cf. Rabinowicz & Rønnow-Rasmussen, 2006, p. 120).
Return now to the previously mentioned counterexamples which were meant to illustrate the wrong-kind of reasons problem. When the Evil Demon forces us to admire the saucer of mud for its own sake, we are given reasons to do so since we want to avoid torture and death. According to the idea currently under consideration, the reasons in question are not of the right kind, because the property of the saucer of mud being such that favoring it for its own sake can help us avoid torture and death is not actually what makes it good for its own sake. When the Evil Demon forces us to admire him for his own sake, we are given reasons to do so since we want to avoid torture and death. The reasons in question are once again not of the right kind, because the property of the Evil Demon being such that favoring him for his own sake can help us avoid torture and death is not actually what makes him good for its own sake. Consider also:

\[ F7: \text{x is finally better than y (y is finally worse than x)} =_{df} \text{it is fitting to prefer x over y for its own sake so that i) the fittingness of the preference is grounded in factors that make x better than y (y worse than x) for its own sake, ii) the preference represents the factors that ground its fittingness as the things that make x better than y (y worse than x) for its own sake.} \]

\[ F8: \text{x is finally equal to y (y is finally equal to x)} =_{df} \text{it is fitting to be indifferent between x and y, so that i) the fittingness of the indifference is grounded in factors that make x equally good as y (y equally good as x) for its own sake, ii) the indifference represents the factors that ground its fittingness as the things that make x equally as good as y (y equally good as x) for its own sake.} \]

The response currently under consideration is not likely to enthuse a lot of philosophers, since most of them have been taught that conceptual circularity is always to be avoided. However, it is unsurprising to me that there are many philosophically interesting concepts that cannot be analyzed completely independently of one another.\(^46\) Furthermore, it might be that while some cases

\(^46\) Quine (1951) famously suggested that the concept of analyticity cannot be understood without reference to meaning, necessity, and synonymy, all of which appear in its conceptual vicinity. This made many philosophers shy back from the concept of analyticity altogether, but perhaps they were too quick to do so. Dickie (1984) is an early example of a philosopher who embraces conceptual circularity, by suggesting that the concept of an art work cannot be
of conceptual circularity are vicious and ought to be avoided, it seems intuitive to suppose that others are benign and may even be desirable. Discussions in this area have become sophisticated and several different requirements for benignity have been identified and defended by philosophers (e.g. Humberstone, 1997; Keefe, 2002; Burgess, 2008). Here we may content ourselves with supposing a first sign of benignity to be that a circle is large enough to incorporate a rich and illuminating structure, which can also be shown to be theoretically useful in its application. Intuitively, the fitting-attitudes analysis seems as if it is a case in point, even with the abovementioned understanding of the right-kind of reasons. Indeed, at first sight, it seems like a circular version of the fitting-attitudes analysis would be just as rich, illuminating, and theoretically useful as its noncircular counterpart. It is true that the conceptual gain speaking in its favor may no longer be quite as strong, since no strict reduction takes place of value to the presence of reasons for attitudes. We are not brought any closer to the philosopher dream of turning the normative domain into a desert landscape, comprised of only a few fundamental concepts. However, it seems to me that what little is lost in terms of conceptual gain is gained in terms of philosophical understanding. For more on this issue, the reader is referred to the appendix.

On the Normative Property of Being a Reason

There is a meaningful and philosophically important distinction between reasons, on the one hand, and the normative property of being a reason, on the other hand. Reasons are simply facts, understood here as obtaining states of affairs.\textsuperscript{47} The normative property of being a reason is a feature that is meant to

\textsuperscript{47} It should be mentioned that this is very stipulative on my part and that the understanding of facts set out here may not reflect the commonplace within philosophy. For example, Zimmerman (2001, pp. 46–52) worries that facts have different modal properties from states of affairs, and so claims that facts should instead be understood as the obtainings of states of affairs. I find this maneuver troubling for the simple reason that it robs facts of their apparent propositional character. Obtaining states of affairs are propositional in a way that the obtainings of states of affairs are not. This means that if Zimmerman is right about facts having different modal properties than states of affairs, we are faced with a dilemma. Two of the more obvious ways of accounting for the facts that constitute reasons fail to capture either their modal properties or their propositional character. I must set aside this dilemma for now. I will also set aside any worries there might be regarding the ontological status of facts as such. There are philosophers who would deny that there are facts, at least not when these are taken to be something over and above the objects, properties, and relations that constitute them.
be shared by all facts that are reasons. Some philosophers claim that reasons are not part of the furniture of the world. Yet few of them are skeptics about the existence of facts. The question is how this could be: Is it possible to be a skeptic about reasons without being a skeptic about the facts that constitute them? What else could philosophers possibly be referring to when they claim that reasons do not exist? Upon reflection, it becomes clear that what is particularly strange about reasons is not their status as facts, but rather the normative property that the facts that constitute them share, namely the property of being a reason. When a philosopher claims that reasons are not part of the furniture of the world, then, she need not deny that there are facts. Instead, she might claim that the property of being a reason is one of those normative properties that are never really instantiated in any of the states of affairs that obtain.

Given that the focus in this work is primarily on final value rather than reasons themselves the question invites itself why any of this is important. However, if we assume that final value can be understood in terms of the presence of reasons, then it is obvious that coming to understand the latter might be of help in coming to understand the former. This is all the motivation we should need. Another reason for paying attention to the distinction between reasons and the normative property of being a reason is that most normative properties should be assumed to be dependent features. Whenever a fact is judged to be a reason, then, we are committed to thinking that the fact has other properties that can explain, in both a broad and narrow sense, why this is the case.48 Some of the chapters in this work are dedicated to showing that this observation can help us gain a significantly better grip on the concept of final value. In particular, it will turn out that the key to understanding certain aspects of the relations of final value is found among issues to do with the dependence and resultant of the normative property of being a reason. For the sake of preparation, then, we should consider briefly what more we can possibly say about the property in question.

It should be obvious that the normative property of being a reason is a relational one. How the property is relational is revealed when we observe that

---

48 Again, what we are committed to saying is not that the fact constituting the reason is a fact as a result of something else. This may well be true and complicates the overall picture significantly, but what we are committed to saying is merely that the fact is afforded its normative status as a result of something else.
whenever there is a reason, it is always a reason for some agent to respond in certain ways. In other words, when a fact constitutes a reason, it is always a reason for an agent to commit an action, believe something, or direct an attitude into the world. In negative terms, there can be no reasons that are not reasons for someone and there can be no reasons that are not reasons to respond in certain ways. This three-placed view of the relational property of being a reason represents a sort of minimal understanding. This minimal understanding is shared by all reasonable accounts of reasons and is compatible with all the views which will be defended in this work. How the minimal understanding ought to be expanded is a controversial issue that we cannot consider here. However, it is worth bearing in mind that there are significant structural problems that crop up as soon as weight is put behind the distinction between reasons and the normative property of being a reason, especially regarding how it should be made to capture the phenomenon of normative strength.

Finally, I should say something more about the kinds of reasons that will be of relevance here and how they should be understood. The fittingness appealed to by the fitting-attitudes analysis should be interpreted in terms of justificatory reasons. Justificatory reasons are commonly distinguished from explanatory and motivational reasons (e.g. Raz 1975; Smith 1994; Parfit 1997; Dancy 1995). Explanatory reasons are typically the kinds of reasons that are invoked in causal explanations. When a flower withers and dies and the reason for this is claimed

49 As such it is also subject to expansion by philosophers who suspect that it might be a bit too minimal. Skorupski (2010) is among them. He suggests that when a fact is a reason, it is always a reason of a certain degree for an agent to respond in certain ways. Scanlon (2014) also offers up an expansion like this, but he does so in a different way. He suggests that when a fact is a reason, it is always a reason for an agent to respond in certain ways in a certain fixed circumstance. Obviously, the main difference between the two philosophers has to do with their respective views on normative strength. Whereas Skorupski develops an account that makes a direct reference to degree of strength, Scanlon instead relies upon a notion of circumstance. This notion of circumstance is in certain respects vague, but it seems to incorporate most, or many, of the factors that might affect the normative property of being a reason. Among other things, then, Scanlon’s notion of circumstances includes background conditions as well as the presence of other reasons (2014, pp. 30–31). Scanlon thereby hopes that he can account for degree of strength indirectly, in a manner that does not make it into an essential aspect of the normative property of being a reason. On his view, it is only sensible to think and talk about degree of strength if we are considering one reason in relation to other reasons (2014, p. 108). Again, this is an issue about which I will remain noncommittal.

50 Tradition holds that a reason is internal if, and only if, it is constituted by facts about an agent’s motivational attitudes. Otherwise the reason is external. The separation between reasons and the normative property of being a reason suggests an additional distinction: We might say that a reason is intrinsic if, and only if, it is constituted by a fact which owes its normative status to something within the fact itself. Otherwise the reason is extrinsic.
to be its having been exposed to too much sunlight, the reason in question is of the explanatory kind. Motivational reasons are the reasons we take to have motivated responses. When a person shoplifts a loaf of bread from a vendor and claims that the reason for this is that she was terribly hungry, the reason in question is of the motivational kind.\textsuperscript{51} Justificatory reasons are not necessarily invoked when we aim to explain or motivate responses, but are, as the name implies, invoked to justify them. In other words, justificatory reasons are facts that counts in favor of responses in a more robustly normative sense. The justificatory reasons that are of relevance here need not be moral and they need not be strong enough to amount to any other sort rational requirement. I take it that when an object has final value, this does not necessarily entail that we ought to respond to the object in any specific manner or that we are obliged to do so.\textsuperscript{52}

Another way of making the points that have just been made is to say that the reasons that are of relevance to the fitting-attitudes analysis, and hence most of the discussions to follow in this book, are meant to be both justificatory and of the \textit{pro-tanto} variety. This gives us a very general sense of the kinds of things we will be interested in, but admittedly it remains somewhat unclear exactly what it is for a fact to have the normative property of being a reason: What does it mean for something to recommend or speak in favor of a certain response? If we analyze final value in terms of the presence of reasons, then are we not also required to develop an informative analysis of what it means to be a reason? It is commonplace among philosophers to assume that what it means for a fact to recommend or speak in favor of a certain response is not amenable to analysis. The normative property of being a reason is, in this sense, taken to be a primitive one. I cannot provide an analysis of what it means for something to be a reason here and so, within the context of this work, this phenomenon must here be treated \textit{as if} it were a primitive one.

\textsuperscript{51} On the stipulative understanding used here, motivational reasons are explanatory reasons that are applied to account for human behavior. There are other understandings of motivational reasons, which, for example, take them to involve the facts that agents themselves would point to when motivating their behavior.

\textsuperscript{52} In fact, when all other relevant justificatory reasons have been identified and weighed in the appropriate way, it may turn out that even though a person has reasons to respond in a certain way, it is nonetheless something that, on balance, the person ought not to. The person might not even be rationally required to do so, on my understanding of the relevant concepts. This understanding is, however, controversial. For further discussions about the connection of reasons and oughts, see e.g. Raz (1990), Bedke (2009), Skorupski (2010), Hory (2012), Broome (2015), and Parfit (2011).
A Brief Restatement of the Project

To better understand what it means for objects to have final value, we need to get a handle on the relations that it stands in to other notions in value theory. This includes other value concepts and other normative concepts more generally. In addition, we need to get a better grip on the dependence of final value upon other things, which, as we have just seen, requires taking a closer look at several distinct phenomena: This includes the notion of evaluative resultance, which involves the relation between final value and the things in virtue of which, or as a result of which, objects are good or bad for their own sakes; this also includes the similar notions of background conditions and constitutive grounds, which involve the relation between final value and the factors that explain why other factors take part in evaluative resultance bases. The following work is meant to help us approach this better understanding and to tighten our grip on some of the second-order issues surrounding final value. Its success is determined by the extent to which it manages to deal with some of those issues in a satisfactory way, while perhaps pointing out some new ones.
Chapter 2
Challenges from Attributivity

The concept of final value is among the most prominent concepts to figure in value theory. As such it has also been the subject of much criticism from philosophers. The following chapter focuses on one influential objection, the purpose of which is to cast doubt over the very coherency of final value. The objection states that to have value is necessarily to be good or bad in a certain way. Geach (1956), who originated the objection, took this to mean that what has value must be good or bad as a kind of thing. His contemporary supporters, among them being Foot (1985, 2001), Hursthouse (1999), and Thomson (1994, 2008), tend to be more liberal in this regard. They say that in addition to having value as a kind of thing, objects may also be good or bad for someone, or to do things with.53 The chapter spells out some of the arguments behind the objection and develops two distinct lines of response: The first line of response

53 On a relaxed understanding, objects may be judged to have value in a way in the sense that some of their properties make them good or bad. In other words, a painting may be judged to be beautiful in a way and this might mean nothing more than that the painting is beautiful with respect to its colors and shapes. On another relaxed understanding, objects may be judged to have value in a way in the sense that some of their properties are themselves good and bad. Accordingly, a painting may be judged to be beautiful in a way and this means nothing more than that the colors and shapes of the painting are themselves the bearers of the relevant value. As we proceed through the discussions to follow, it is important that we bracket these more relaxed understandings of value in a way. Otherwise we risk wasting time by talking past Geach and his followers. We shall soon see that to capture the family of concepts that they have in mind, it is not enough that we point out what objects have value or perhaps even what makes said objects good and bad.
attempts to show that the objection fails to establish its case against the coherency of final value. Instead, the objection manages to highlight that there may be ambiguities involved in the use of value concepts. Consequently, we need to be very careful when we invoke such concepts in ethical theorizing, and while formulating arguments and inferences.

The second line of response attempts to show that even if the objection did manage to establish its conclusion, the more salient features of the concept of final value could still be salvaged by being couched in slightly different terms. The idea is that objects can be good or bad in a certain way, but in such a manner that it also becomes fitting to value the objects for their own sake. This response to the objection is the more interesting one for two main reasons: The first reason is that it would indicate that final value is not just one of the most prominent concepts in value theory, but it is also one of the most fundamental. By this I mean that many types of value can come in the final variety, including those that are emphasized by Geach and his followers. The second reason why this is the more interesting response is that it contributes to the main endeavor undertaken throughout this work, which is to clarify the dependencies and relations that final value stands in to other things. More specifically, it would highlight some important and often overlooked aspects of the relation between the concept of final value and value in a way.  

Before we move forward a couple of clarifications need to be made, the first of which concerns the precise understanding of value in a way. Hereafter, we should interpret this as a rather technical notion, covering several different value types that are represented in our evaluative talk and practices. For the sake of brevity and simplicity, two such concepts will be highlighted and treated as paradigmatic examples of value in a way. The first is the previously mentioned concept of kind-value, which denotes the type of value that accrues to objects
that are good or bad as the kinds of things that they are. When an object is judged to be a good toaster, this is the value concept that we have in mind. The second is the previously mentioned concept of personal value, which is the type of value that accrues to objects that are good or bad for someone. When we judge prescription glasses to be good for someone with bad but correctible eyesight, this is the concept that we have in mind. There may be other types of value that should also be gathered under the banner of value in a way, but to discuss all of them in detail would not be possible in the present context. Instead, my assumption is that whatever observations we make about kind-value and personal value will also hold true of value in a way more generally.

The second clarification I wish to make concerns the intended targets of the objection by Geach and his followers. In their original discussions of value in a way, there are few explicit mentions of final value to be found. Instead, their arguments tend to be couched in terms of the objects being good or bad period. This could be taken to suggest that their discussions about value in a way are not meant to be of relevance to the concept of final value in the first place. There are several reasons why we should be skeptical of this way of seeing things. The main one is the fact that proponents of the concept of final value have often talked about this type of value in terms of objects being good or bad period. This includes G. E. Moore, who also happens to be a frequent target of Geach and his followers. Another target, mentioned throughout the discussions of both Foot and Thomson, are utilitarians that ascribe final value to abstract objects like facts and events. If it is true that both Moore and the utilitarians can escape the objection under consideration simply by clarifying that what they are interested in is final value, then one begins to wonder about the incredible influence that Geach and his followers have had on discussions within this field.

55 Thomson (2008) is among the philosophers who are most eager to point the many varied ways of being valuable that Geach misses in his original discussions. Besides saying of something like a book that it is good as a piece of literature, as a source of information, or as a drink coaster, she thinks that it may also be good for someone, good for certain purposes, and good to read. Again, we shall have to leave most of these concepts unexplored here, although it is my suspicion that some of these additional examples of value in a way are reducible to the concepts of kind-value or personal value. See also Skorupski (2010, p. 86).

56 Obviously, the fundamentality of final value cannot be defended with respect to all kinds of values in the following work. A selection is made based on what seems salient to contemporary discussions in value theory. It might be wondered, for example, whether value concepts like admirability and respectability can come in the final variety. Can we say of a person that she is admirable or respectable for her own sake? We will have to leave such questions unexplored in what follows and acknowledge the possibility that although the concept of final value might be among the most fundamental concepts to be found in value theory, it might not cut across the entire evaluative domain. This, at the very least, remains to be shown.
Their arguments in favor of value in a way will then start to look a bit trivial, as if they were mostly engaged in philosophical shadowboxing. To avoid this uncharitable interpretation, I shall hereafter assume, as many philosophers have done before me, that the objection is indeed meant to be relevant to the concept of final value as well.57

The Case against Final Value

Recall the friend who draws our attention to a couple of paintings and says of one of them that it is good for its own sake. This provokes our interest and we proceed to ask her why the painting is good for its own sake. Her answer is that it just is. This makes us feel a bit perplexed. As we have seen, the use of value concepts must reflect the fact that if objects have value, then this has something to do with what the objects are otherwise like. Objects are always valuable in virtue of, or as a result of, something. Suppose therefore that our friend elaborates on her original answer and explains that the painting really has value because it features certain colors and shapes. She persists in her claim, however, that the painting is only good for its own sake. At this juncture, we may well disagree that final value ever depends upon colors and shapes in the way she claims. Nevertheless, we would be hard pressed to say that her explanation betrays any conceptual confusion. Once she has explained why the painting is good for its own sake, the initial feeling of perplexity that we might have had should be largely allayed. Yet, some philosophers argue that this is not quite enough. They maintain that uses of value concepts must accommodate yet another second-order requirement, namely one which states that objects are always good or bad in a certain way.

There are plenty of examples that tell in favor of such a conceptual constraint. If the object under consideration is a painting, then it would make sense to say that in addition to being good because it features certain colors and shapes, the painting is good as a painting, as piece of art, as part of a decor, or as a representative of a cultural movement. It would be more than a little peculiar if our friend claimed that the painting is good but denied that it had value in any such sense. To my knowledge, the first philosopher to ever explicitly endorse

57 For a defense of Moore that to a large degree is about these kinds of exegetical and interpretive issues, see Tucker (2017). He argues that throughout much of her discussion, Thomson does indeed engage in a bit of shadowboxing, at least in regard to her criticism of Moore.
and elaborate upon a conceptual constraint along these lines is the previously mentioned Geach (1956). He famously argues that evaluative terms may be used grammatically in both an attributive way (as in the claim ‘this is a good book’) and a predicative way (as in the claim ‘this book is good’), but from a conceptual standpoint they are always used attributively.\textsuperscript{58} In order to defend and clarify this view he seems to me to primarily develop two distinct lines of argument: The most discussed argument proceeds from the notion that value concepts are revealed to be attributive when we look to how they tend to figure in arguments and inferences.\textsuperscript{59} For the sake of illustration, let us begin by considering the following examples:

\textit{Bottom-up Red:}

\begin{itemize}
  \item x is red
  \item x is a car
\end{itemize}

\begin{itemize}
  \item Therefore: x is a red car
\end{itemize}

\textit{Top-down Red:}

\begin{itemize}
  \item x is a red car
\end{itemize}

\begin{itemize}
  \item Therefore: x is red
  \item Therefore: x is a car
\end{itemize}

\textsuperscript{58} The same line of reasoning is of course meant to apply to comparatives such as ‘better’, ‘worse’, ‘equally as good as’, and so on. Since one of the paintings mentioned in the original example is meant to be better than the other, this must be because the painting in question is better as a piece of art, as part of a decor, or as a representative of a cultural movement, or something similar. On the view of Geach and his followers, it would not make any sense if our friend claimed that one painting is better than the other, but without being better in any similar sense. In other words, whenever one object is better, worse, or equally as good as another object, the former will always be better, worse, or equally as good in a way.

\textsuperscript{59} Geach first hints at the test by his very definition of the attributive. He states that “in a phrase ‘an A B’ (‘A’ being an adjective and ‘B’ being a noun) ‘A’ is a (logically) predicative adjective if the predication ‘is an A B’ splits up logically into a pair of predications ‘is a B’ and ‘is A’” (1956, p. 33). Now, it should here be mentioned that Geach says that his arguments are about terms and does not address abstract concepts. However, as Orsi (2015, p. 48) notes, the very same arguments should apply to concepts as well. In this case, the semantics and the conceptual seem to go hand in hand.
**Substitution Red:**

x is a red car
x is a status symbol

Therefore: x is a red status symbol

These arguments and inferences involving the concept of being red seem to be valid. If the premises are true, then the conclusions must be true as well. When taken together, they indicate that the concept of being red can be used rather freely across different formal and linguistic contexts. Interestingly, this does not appear to be the case when it comes to the terms involved in the following inferences:

**Bottom-up Big:**

x is big
x is a flea

Therefore: x is a big flea

**Top-down Big:**

x is a big flea

Therefore: x is big
Therefore: x is a flea

**Substitution Big:**

x is a big flea
x is an animal

Therefore: x is a big animal

These lines of inference involving the concept of being big are not obviously valid. Even if the premises are true, there is no guarantee that the conclusions must be true as well. This seems to show that the concept of being big is not
quite as free-spirited as the concept of being red is. Compare the examples given so far with the following:

**Bottom-up Good:**

\[
\begin{align*}
\text{x is good} \\
\text{x is a father} \\
\hline
\text{Therefore: x is a good father}
\end{align*}
\]

**Top-down Good:**

\[
\begin{align*}
\text{x is a good father} \\
\hline
\text{Therefore: x is good} \\
\text{Therefore: x is a father}
\end{align*}
\]

**Substitution Good:**

\[
\begin{align*}
\text{x is a good father} \\
\text{x is a bowler} \\
\hline
\text{Therefore: x is a good bowler}
\end{align*}
\]

These lines of inference involving the concept of being good are not obviously valid. If the premises are true, then there is no guarantee that the conclusions must be true as well. This indicates that the concept of being good cannot be used freely across different formal and linguistic contexts. In this sense, the concept of being good seems to have more in common with the concept of being big than it does with the concept of being red. The crucial question is of course what conclusions we should draw from these initial observations. Heed needs to be taken of the fact that there are at least two distinct ways of capturing these patterns of validity. The explanation favored by Geach himself is a monist one stating that only attributive uses of value concepts are meaningful. This explanation maintains that while the concept of being red is essentially predicative, the concepts of being big and good are essentially attributive. This means that unlike the concept of being red, the concepts of being big and good
are in a sense fettered to categories of object. The concept of being big is understood differently depending on whether we are talking about cathedrals specifically or buildings in general, just as the concept of being good is understood differently depending on whether we are talking about fathers or bowlers. The monist explanation faces several difficulties, among which is the fact that it seems to lack a strong philosophical rationale. The inferences and arguments just provided make it very clear that value concepts are sometimes used attributively, but what reasons are there to suppose that they cannot be used predicatively as well? Why should we not suppose that value concepts can actually come in both the attributive and the predicative variety?

One alternative to the explanation favored by Geach is the pluralist one. This explanation maintains that value concepts can come in both the attributive and predicative variety. So, while the concept of being red is essentially predicative and the concept of being big is essentially attributive, the concept of being good features an ambiguity in this regard. This line of reasoning opens the door for an additional response to the first argument offered by Geach. The response says that the inferences and arguments exemplified above do not appear so strange because value concepts cannot be understood without reference to kinds. Instead, perhaps the inferences and arguments appear so strange to us because they involve a potential ambiguity about the type of value that is at stake. In ordinary life, contextual clues will often give us enough information to determine whether an attributive or predicative use is relevant in a specific linguistic context, but such clues are obviously missing from the examples given here. This entails that once the assumed ambiguity is removed from the inferences and arguments involving the concept of being good, the patterns of

---

60 Geach and his followers are more radical than this, because they think that the associated kinds are meant to be thick enough to set the standards for bigness or for goodness. We return to this point later. Geach (1956, pp. 33–34) originally states things in more linguistic terms, by saying that the adjectives ‘big’ and ‘good’ cannot be used without an associated noun. Thus understood, the monist explanation tempts us toward the suggestion that value terms behave like two-place predicates, taking as their arguments some object on the one hand and some kind on the other. Although this is right to some extent, it would perhaps be more appropriate to liken value terms to predicate modifiers (cf. Rind and Tillinghast, 2008, pp. 83–85).

61 This kind of pluralism has been explicitly endorsed by Hare (1957, pp. 107–108), Pigden (1990, pp. 131–132), Zimmerman (2001, pp. 20–25), Sinnott-Armstrong (2003), Arneson (2010), Sturgeon (2010), Smith (2010), Kraut (2011, p. 179), Orsi (2015, pp. 58–61), Rowland (2016a, 2016b), and Tucker (2017). These philosophers are pluralists and so admit that value has both an attributive and predicative use. This is not just true in a grammatical sense, but in a deeper conceptual sense as well. However, I am less than confident that they would all agree with the deeper pluralist explanation that is about to be presented as well. On that and any other related exegetical issue that might be raised here, I wish to remain silent.
validity should shift accordingly. In other words, the result of such a procedure should be that we can state, with a much higher degree of confidence, that the inferences and arguments are in fact valid. Let us next consider the following:

*Final Bottom-up Good:*

\[
\begin{align*}
x \text{ is (finally) good} \\
x \text{ is a father} \\
\phantom{x} \\
\therefore x \text{ is a (finally) good father}\end{align*}
\]

*Final Top-down Good:*

\[
\begin{align*}
x \text{ is a (finally) good father} \\
\phantom{x} \\
\therefore x \text{ is (finally) good} \\
\therefore x \text{ is a father}
\end{align*}
\]

*Final Substitution Good:*

\[
\begin{align*}
x \text{ is a (finally) good father} \\
x \text{ is a bowler} \\
\phantom{x} \\
\therefore x \text{ is a (finally) good bowler}
\end{align*}
\]

By introducing the parentheses as a clarification, it is here made explicit that what we should have in mind is specifically some sort of final value. Not surprisingly, this makes the inferences and arguments appear valid. If the premises are true, then the conclusions must be true as well. In most cases, the

---

62 There are a couple of readings of this claim that we should try to set aside here. On one such reading, x is good for its own sake as a father. The possibility of this kind of reading shall be defended later in the chapter. On another reading to be avoided, x is good for its own sake and x is good as a father. Here, the claim should just be understood as saying that x is good for its own sake and x is also a father. When understood this way, the conclusion follows.

63 A similar sort of reconstruction seems to be endorsed at one point by Zimmerman (2001, p. 21), although he does not consider inferences of all three types and prefers to couch his reconstructions in terms of intrinsic value rather than final value. For this reason, Zimmerman can perhaps be regarded as being a pluralist about the use of value concepts as well as a subscriber to the deeper pluralist explanation that has just been offered.
context makes this final reading of the relevant inferences and arguments less inviting, and so the possibility that there is an ambiguity here can easily escape attention. In reality, though, we can very easily express value judgments and intend for the final reading to be the correct one, even though we do not explicitly mention the modifiers ‘final’ or ‘finally’. If we are given enough leeway, we can repeat this clarificatory procedure and apply it to kind-value and personal value as well, although the result is likely to be much more cumbersome and for that reason might appear less convincing. The upshot of Geach’s argument, then, is at most that we need to be very careful when we make inferences with ambiguous concepts. Perhaps it also reminds us that just because an object is good as one kind of thing, it does not necessarily follow that the object is good as another kind of thing. It also shows that just because an object is good for its own sake, it does not necessarily follow that the object must also be good as a kind of thing, nor vice versa. For proponents of the idea that there are both predicative and attributive uses of value concepts, none of this should seem very surprising. At this juncture the objection might be made that in everyday life we rarely encounter any troubles with understanding the meaning of value judgments. There just do not seem to be any major

---

64 Almotahari & Hosein (2016) have recently offered a defense of Geach that touches upon some of the issues that have been raised so far in this discussion. The most contentious part of their defense can be understood as relying upon the principle that if there are predicative uses of value concepts, then adjectives such as ‘good’ should be able to figure predicatively in linguistic contexts such as ‘X is a good father’. However, they maintain that this is in fact not possible. The general principle upon which they rely has since been shown to be false by Byrne (2016). He objects to the principle because it gets the wrong results when we look to examples outside of the evaluative domain. We need not go into detail about this objection now, for it seems to me that regardless of whether Byrne is right about this, there might still be room for ambiguities here. In other words, the fact that the adjective ‘good’ can figure predicatively in linguistic contexts such as ‘X is a good father’ is supported by intuitions that are independent of the principle he objects to. When Dickens wrote, “it was the best of times, it was the worst of times”, he did not necessarily commit himself to an attributive use which relates the relative comparatives to the kind times. Conversely, when the Bible says that “God saw all that he had made, and it was very good”, the deity is not necessarily ascribed a predicative use that leaves the relevant adjective disconnected from the kind world. The intuition that there is room for ambiguities in cases like these can of course be challenged, but to simply reject them would be tantamount to begging the question against the pluralist explanation that has just been given. In fact, when Almotahari & Hosein (2016, pp. 1493–1495) do try to offer up reasons to reject the intuition, they come close to begging the question anyway. They deny the possibility that there might be ambiguities here precisely because it makes seemingly invalid inferences along the lines that we have looked at above subject to valid readings. Clearly, this is a point at which the discussion risks devolving into mere table-thumping over intuitions, because I just do not share theirs. See Peace (2017) for some additional criticism of the test by Almotahari & Hosein and for alternative considerations against the possibility that objects can be plain good.
ambiguities in this regard that we need to be especially careful of. Proponents of the pluralist explanation should here respond by once again putting emphasis on what is common and expected in certain linguistic contexts.

In narrow contexts, where attributive uses of value concepts are common and expected, we can often take it for granted that value judgments should be understood attributively. If someone were to say, within the context of a watching a dressage competition, that Valegro was a good horse, there would be no need to consider predicative uses of the relevant value concepts. However, if the misty-eyed owner of Valegro were to say, long after the horse has been put out to pasture, that Valegro was a good horse, there may be room for ambiguity as to what is really meant. It also needs to be recognized that there are categories of object that are, as a matter of norm, never directly referenced by value judgments unless the value judgments are to be understood in an attributive manner. If a person were to say that she has a good toaster, there is also no need to consider predicative uses. This does not mean that the norm in question is a formal one in the sense suggested by Geach and his followers, but rather something to do with the pragmatics of our evaluative talk and practices. Indeed, there may in fact be categories of object that work in the opposite way, in that they are referenced by value judgments only in so far as the value judgments are to be understood in a predicative manner. If a person were to say of her mother that she was a good woman, she should probably not be understood as saying that her mother was good as far as women are concerned.

This line of reasoning reminds us of the fact that the pluralist explanation is given additional weight by substantive examples. The history of moral philosophy is riddled with first-order theories that ascribe final value to objects, without relating them to kinds in the way that is required by the monist explanation. There may be something strange in the claim that a painting is just good for its own sake, but as we have just seen this might not be because this claim is meaningless. What makes this claim strange is rather that it is obviously false or perhaps that it flies in the face of our linguistic expectations in more nuanced ways than what Geach has in mind. The strongest support for the monist explanation seems to be the initial feeling of perplexity that some of us share upon hearing it said that certain objects is simply good or bad for their own sake. However, for a proponent of pluralism it will seem very reasonable to describe this perplexity as the result of a first-order disagreement. Kraut (2011, pp. 173–175) points out that the perplexity may be caused by our thinking that such claims are obviously false, while Rowland (2016a, p. 1377) describes cases where it may result from our thinking that they are obviously true. If this is right, then the norms mentioned in the previous paragraph may have linguistic dimensions, even though they are not formal in the strict sense that Geach and his
peculiar when we imagine that the object under consideration is not a manmade artifact, but something like happiness, preference satisfaction, the experience of pleasure, or the good will. There is no obvious conceptual confusion in judging any of these objects to be good for their own sakes.\(^66\) In the context of giving substantive examples like these ones, it is also worth considering ascriptions of final value to very abstract kinds. It is not particularly hard to see how the value of concrete objects could be related to kinds in the way that is required by the monist explanation, but the question is what we should say of the value of more abstract categories of object. If we are to deny that all ascriptions of this sort are coherent, we need stronger reasons than we have been given so far.

On a related note, it should be mentioned that some philosophers have argued that when we relate value to kinds, there is no formal constraint governing how abstract those kinds may be. The philosophers in question even suggest that it is against the background of abstract kinds that apparently predicative uses of value concepts are given meaning. Mackie seems to come very close to endorsing this idea when he states that an object is just good if it is “a good thing to happen” or “a welcome occurrence” (1977, p. 57).\(^67\) In his original discussions, Geach tries to block this way of thinking by maintaining that value must not be related to kinds that are too descriptively vacuous. As he sees things, the relevant kinds must be able to provide us with an adequate “criterion of identity or standard of goodness” (1956, pp. 40–41). Thomson agrees and says that “if what being a K is doesn't itself set the standards that a K has to meet if it is to be good \textit{qua} K, then nothing sets them—thus there is no

\(^{66}\) It can easily happen, and it is well worth remembering while reading this work, that some of our first-order judgments about value become so ingrained that we mistakenly regard them as indicative of something stronger, like a constraint governing the proper use of value concepts. Geach and his followers tend to be proponents of first-order theories that are highly influenced by Aristotelian ethics. Just as we can understand what a good toaster is by considering the general function and uses of toasters, the hope is that we can understand what a good human being is once we come to grips with what it means to be a human being. From this perspective, it is perhaps not surprising that they endorse a conceptual constraint barring predicative uses of value concepts. An obvious downside of this is that their suggestion renders some of the central insights from Aristotelian ethics trivial, when in fact they are not. If it is true that good human beings and the good lives of human beings cannot be understood without referring to what characterizes the human kind, this is a more interesting truth than Geach gives it credit for.

\(^{67}\) Mackie takes this to indicate that Geach was wrong in his contention that all uses of the term ‘good’ are connected to some determinate noun. This may seem strange, since the term ‘occurrence’ is a noun. Mackie should presumably be understood as having in mind nouns that pick out objects with a determinate function or goal—in other words, objects belonging to kinds that on their own manage to provide a criterion of identity and standards of goodness.
such property as being good *qua* K” (2008, p. 21). Neither Geach nor Thomson provides their claims on this point with much argumentative support. However, the temptation to use this disadvantage to gain further leverage against their views should on my view be avoided, as it is unnecessary.

To suggest that for an object to be just good is for it to be good relative to a very abstract kind, like being good as far as facts are concerned, is effectively to agree with the first argument made by Geach. It would be the same as saying that predicative uses of value concepts are only seemingly predicative and, as we have just seen, we have not been given sufficiently strong reasons to think that this is the case. Both the monist and the pluralist explanation appear to be equally capable of capturing the patterns of validity that were uncovered earlier. The pluralist explanation is also given further support by considering what happens once the assumed ambiguities are removed from the inferences in question. This fact, together with the appeal of substantive examples and the possibility of ascribing value to abstract kinds, makes it seem as if the monist explanation needs additional defense. As a matter of fact, Geach (1956, p. 35) does formulate a second argument in favor of his view which is meant to warn us away from the possibility of pluralism. This argument states that if it is true that value concepts are amenable to both predicative and attributive uses, then there is a risk that these uses would turn out to represent completely distinct concepts with little to no connecting tissue. There should be some conceptual resemblance between the uses that can help explain the intuition that both can come to involve genuinely evaluative concepts. This is clearly an interesting and worrying argument that merits closer consideration.

Now, it is often clear from their discussions that the primary targets of Geach and Thomson are philosophers belonging to the tradition of Moore and Ross (cf. Geach, 1956, pp. 35–36; Thomson, 2008, p. 14). The reason is presumably

68. It would obviously be no argument to suggest that a lack of such a criterion or standard of goodness makes much more difficult the first-order endeavor of establishing what objects have value. This may well be true, especially if we couch our first-order theories within Aristotelian ethics, but this would say very little about how value concepts may actually be used.

69. Compare this to the arguments of Peace (2017), according to which we should turn our focus from the concept of being just plain good to the concept of being good as a state of affairs. This presumes that the two are in fact very different from each other. I think this is right.

70. Geach (1956, p. 35) understands the point semantically and worries that predicative and attributive value terms might then turn out to be mere *homonyms* (cf. Pigden, 1990, p. 133). This is indeed a cause for concern, but it is unfortunate that Geach can only think to mitigate the worry by outright denying the meaningfulness of one of the two terms. We shall soon explore the more promising strategy on offer that does not throw the baby out with the bathwater in this way.
that the latter philosophers support views that seem to afford an additional air of mystery to predicative uses of value concepts: One such view is that value concepts are primitive and so not amenable to conceptual analysis. Another such view is that value concepts pick out genuine properties or relations that cannot be accurately described in a wholly nonnormative language. According to the assumption made in this work, for example, value concepts can be analyzed in terms of fitting attitudes. This includes the concept of final value, which is here understood in terms of there being reasons to either favor or disfavor objects for their own sake.\textsuperscript{71} Neither Geach nor Thomson would presumably question the intelligibility of the idea that there can be reasons of this sort and so their worry about the property of being just good should be largely mitigated—barring, of course, any skepticism they might have with regard to the fitting-attitudes analysis as such.\textsuperscript{72} Indeed, Thomson herself comes close to acknowledging this way of thinking when she writes:

Thus when they ask, ‘is pleasure good?’ what they are asking is just, simply, whether other things being equal, we ought to promote pleasure. But if that is what a philosopher means when he asks, ‘is pleasure good?’ we might well wonder why he asks his question in those words. Why does he help himself to the word ‘good’? Why doesn’t he just, simply, ask whether other things being equal, we ought to promote pleasure? (2008, pp. 15–16)

\textsuperscript{71} In his response to the arguments by Geach and Thomson, Zimmerman (2001, pp. 20–25) endorses a similar line of reasoning. He claims that being intrinsically valuable is the same as being good or bad in a moral way. By this he seems to mean that intrinsic value generally involves a moral requirement for attitudes. Although this is a sympathetic view in some respects, it is not clear to me why final value would be generally accompanied with moral reasons. The issue is a dicey one because no philosopher has yet managed to draw a line between reasons that are moral and those that are not. Still, it would be somewhat surprising if it turned out that whenever we judge objects to be good or bad for their own sakes, we are also judging that there are demands of a distinctly moral character present.

\textsuperscript{72} In her discussions, Thomson (2008, p. 14) acknowledges that there are concepts that can be used in both a predicative and an attributive sense. One of the examples that she mentions is the concept of being famous. However, she thinks that there is a crucial difference here in that we have a clear idea of what kind of property is meant to be picked out by the concept in question. The property of being simply famous is identical to the property of just being well known, with which we are familiar. She asks: “What is the property that Ross claims is ascribed to knowledge, or pleasure, by a philosopher who says ‘That is good’ of it?” (2008, p. 14). I cannot presume to speak for Ross here, but if the question is meant to be directed to proponents of final value in general, then it seems as if we have a ready answer. As we have just been reminded, the fitting-attitudes analysis states that the property of being good is to be the fitting object of a certain kind of positive response.
Thomson could perhaps be taken to assume that the only answer to the question she poses in this paragraph is that the property of pleasure being good is also the property in virtue of which, or as a result of which, we ought to promote pleasure. If she is right about this, then the challenge set out by Geach seems largely untouched. For it will then remain unclear that it is meaningful to ever speak of the property of being just good. However, this is obviously not the only answer available to us. The philosopher imagined by Thomson might follow the line suggested in this work and insist that for pleasure to be just good is for it to be the case that we ought to promote pleasure—or, more modestly, for pleasure to be just good is equivalent to there being reasons to favor pleasure. This is not just meant to capture what the philosopher means, moreover, but what we all commit ourselves to when we judge objects to be good or bad for their own sake.73 Herein also lies the key to answering the second argument brought up by Geach. The argument is that if we allow that there may be both predicative and attributive uses of value concepts, these uses might end up having little to no connecting tissue. We are now able to respond to the argument by saying that the connecting tissue that ties the evaluative domain together and unites value concepts of different types is the presence of reasons for attitudes. Throughout the rest of this chapter my aim is to argue in favor of this response.74

The support in question comes in the form of an attempt to analyze value in a way in terms of fitting-attitudes. Again, for reasons of brevity and simplicity, my attention will be on the concepts of having value as a kind of thing and having value for someone. In addition to showing that the concepts are amenable to the fitting-attitudes analysis and so do have a lot of connecting tissue to other types of value, my applications will indicate that value in a way is also, in fact, subject to the distinction between final and instrumental value.75 This means that

73 The suggestion is sometimes made that the concept of having value without being good or bad in a certain way is a technical invention of philosophers which is not really represented in our evaluative talk and practices. When he comments on predicative uses of value concepts, Geach at one point asks: “But how can we be asked to take for granted at the outset that a peculiarly philosophical use of words necessarily means anything at all?” (1956, p. 36). Thomson similarly writes her discussions: “Nevertheless, it is right to say that ‘good’ has a predicative use; but to find it, you have to look in books that say ‘Philosophy’ on the cover” (2003, p. 94).
74 The strategy of using the fitting-attitudes analysis specifically to respond to the skepticism of Geach and to account for the connecting tissue that is supposed to join predicative and attributive uses of value concepts can be traced back to Pigden (1990). The same strategy is more recently discussed and endorsed by Rowland (2016a, 2016b).
75 Recent attempts to analyze value in a way in terms of fitting-attitudes have been made by several philosophers, including Darwall (2002), Rønnow-Rasmussen (2011), Schroeder (2010), Skorupski (2010), Orsi (2015), and Rowland (2016a, 2016b). The argument that the concept of being valuable for someone is subject to the distinction between final and
objects can be finally good as a kind of thing, just as they can be finally good for someone. This point should take care of any remnant that there might be from the semantic and conceptual attack on final value that has been roughly outlined in the preceding discussion. However, before we proceed with any attempts to analyze the concept of being valuable in a way, we should start off by clearing the field of some alternative and very common understandings of the concepts in question. The reason is not necessarily that these common understandings are illegitimate, perhaps, but rather that they do not seem to involve value in the sense that we should be primarily interested in here. As we shall see, these common understandings tend to rely upon the view that value in a way can be explicitly couched in wholly nonnormative terms.

Descriptivism about Value in a Way

I shall soon continue with my defense of the concept of final value and offer further support in favor of its importance n. To reiterate, I will do so by arguing that the distinction between final value and instrumental value cuts across this domain, meaning that even value in a way can come in the final or instrumental variety. Before that, though, I wish first to make some remarks here about the theory of descriptivism. In the introduction to this chapter, it was mentioned that value in a way is here understood as an umbrella term which gathers the concepts of kind-value and personal value. In the introductory chapter, it was also maintained that value is generally situated within a broader normative domain. This was taken to mean that all value, regardless of type, has an intimate connection to reasons for attitudes. It is important that we keep this in mind, because certain analyses of value in a way may suggest themselves that account for the relevant concepts in nonnormative terms—terms that are seemingly divorced from any normative implications. This is precisely the approach that descriptivism advocates for. Consider the following analysis:

\[ D1: \quad x \text{ is a good } K =_{df} x \text{ conforms to the standard that is associated with things that are } K. \]

instrumental value is first explicitly put forward by Rønnow-Rasmussen (2011, pp. 57–60). Unfortunately, I am not aware of any philosopher who agrees with me that the concept of being valuable as a kind of thing can also come in either the final or the instrumental variety.
Since this analysis is meant to be nonnormative, the notion of standard that it relies upon should be understood in deflationary terms. Roughly, for something to be a good toaster is for it to be able to do the things that toasters are designed to do, like warming bread relatively quickly, evenly, and efficiently. Fulfilling the requirements of standards, then, is just a matter of getting enough check marks on some determinate list of nonnormative features. This pattern of analysis seems to me problematic, mainly because it seems to divorce value in a way from any normative implications whatever. Mackie (1977, pp. 25–27) at one point discusses value judgments involving standards and points out that even those of us who would deny that value is a genuine property or relation in the world tend to be quite comfortable in accepting that there are facts about the extent to which objects manage to conform to the requirements of standards. It is obviously an objective matter, he says, whether and to what extent a toaster manages to conform to the requirements of the standard that is conventionally associated with toasters. This is just not the sort of thing that keeps metaphysically frugal philosophers up at night, and which makes the topic of value of such great philosophical interest.

My view is that this should not be taken to mean that when we form judgments based on standards, the results are not genuine value judgments, for this is obviously too strong. Even if we are opposed to the pattern of analysis expressed by D1, we could readily admit that the requirements of standards play an important role for the formation and justification of value judgments. What we should be insisting on is merely that genuine value judgments cannot be wholly understood in terms of judgments about the extent to which objects manage to conform to the requirements of such standards. In other words, an

---

76 For this reason, it may also be tempting to understand the notion of a standard here in prototypical terms (cf. Thomson, 2008, p. 19). A good toaster is an object that is relatively like the prototypical members of the toaster class. A bad toaster is an object that is relatively dissimilar to the prototypical members of the toaster class. There are reasons to be wary about this understanding though. For it is not at all obvious that the standards appealed to in a judgment about something being valuable as a kind of thing can always be derived from observations about what is typical of objects belonging to this kind.

77 Many value judgments are formed against the backdrop of standards. Some of those standards are implicit and personal, while others are explicit and institutionalized. Indeed, even value judgments that do not typically take place against the backdrop of standards may do so at some point in the future. It seems strange to suggest that they would at that point cease to constitute genuine value judgments. When dog shows were invented and there were suddenly specific standards on which to rely to determine what makes a good or bad canine specimen, it is not as if value judgments regarding dogs necessarily ceased to be genuine value judgments.

78 Philosophers often assume that the resultance base for value must have a specific structure. For example, it was mentioned in the introductory chapter that objects were thought to have final
awareness of the standards that characterize certain kinds of object is not equivalent to an understanding of what it means for something to be good or bad, at least not in any philosophically interesting sense. It is also important to note that this objection need not be motivated by a general skepticism toward descriptivism. That is to say, we do not have to rely on the idea, which is often defended by philosophers in these sorts of contexts, that there is an insurmountable conceptual gap between value judgments and nonnormative judgments. Instead, the rejection is motivated by the view that judgments about the extent to which objects manage to conform to the requirements of standards alone do not entail anything about reasons for attitudes.

Of course, it may be objected that even this more modest view begs the question against my opponents, since philosophers like Geach and Thomson are in fact defenders of the kind of descriptivism which tends to understand value in terms of conforming to the requirements of standards. Perhaps there is something to this, but it needs to be remembered on which side the burden of proof currently lies. There may well be instances where the conceptual gap between value judgments and nonnormative judgments is only apparent and not real—where careful reflection reveals that the seemingly open questions that Moore (1993/1903, pp. 67–68) would have asked will turn out to be closed. The suggestion that we have come across such a case in all this talk of conforming to the requirements of standards is, in any event, not very convincing. This is not only true because of such judgments apparently not entailing anything interesting about reasons for attitudes, but also because of the well known difficulties with linking questions about what characterizes kinds and questions about what makes the individual members of these kinds good or bad.79 At this point, we may therefore simply point out that if our understanding

value only in virtue of, or as a result of, their intrinsic properties. It has since become increasingly accepted that the value of objects can result from extrinsic properties as well. If this idea is on the right lines, then it should even be possible for objects to have final value in virtue of, or as a result of, fulfilling the requirements of their standards.

79 The issue is whether everything most of us would regard as good can be associated with any standard in the first place. Many philosophers have questioned this, among them Copp (2011, p. 182), who is especially interested in the concept of a good human being. He denies the view that what it means to be a human being sets the standards that must be adhered to by a person if she is to be good as a human being. He asks us to suppose that two people disagree about whether a given trait, such as kindness toward other human beings, is a virtue. If Geach and Thomson are right, then, this disagreement is a disagreement about what it means for a person to be a human being. Either one of the relevant persons is ignorant about this fact or confused about the meaning of value concepts. Copp insists that this cannot be right.
of kind-value turns out to be too reliant on the plausibility of this kind of reductive analysis, then this makes the concept suspect.

The warning that has just been made regarding the concept of kind-value applies just as well to the concept of personal value. Until very recently, a lot of discussions surrounding this concept have tended to be rather dismissive about its philosophical significance—or even its coherency. This is primarily due to the influence of Moore (1993/1903, p. 150). He famously worries that the concept of being valuable for someone looks too much like the concept of being true or false for someone, which he thinks is relativistic nonsense. The only sense Moore can make of judgments to the effect that one object has value for someone is deflationary: It would just mean that it is good that the object is in the possession, or occurs in the life, of the person in question—or that a good object is in the possession, or occurs in the life, of the person in question. Other dismissive accounts have it that an object has value for someone when the person desires the object or judges the object to have value. Among the slightly more interesting analyses to come out of recent discussions about the concept of personal value are those that rely heavily on the concept of wellbeing:

\[ D2: \quad x \text{ is good for } P =_{df} x \text{ is conducive to the health or wellbeing of } P. \]

Just like the notion of a standard, the concept of wellbeing might be given a normative reading. Another reading, which is the one I wish to warn readers against, is the one which takes wellbeing to be a completely nonnormative concept. For example, one suggestion has it that wellbeing should be accounted for in terms of the functions of objects and what the objects are characteristically designed to do. Roughly, something is good for a toaster when it is conducive to the toaster’s ability to conform to the requirements of its standards. There are other ways of cashing out this general kind of deflationary understanding, but this one might seem relatively plausible, since it wed ds the concept of being valuable for someone to the concept of being valuable as a kind of thing. This is attractive, in part, because questions regarding the functions of objects have an intuitive connection to questions about the kind of object that it is in the first

---

80 See also Hurka (1987, p. 73) for some support of this reading of Moore.

81 There are several ways in which an object can be conducive to the health or wellbeing of a person. For example, this may happen because the object causes an increase in the health or wellbeing of the person or because the object is constitutive of the health or wellbeing of the person. Thanks are owed to Toni Rønnow-Rasmussen for making me aware of this point (in private conversation).
Toasters can warm up bread relatively quickly, evenly, and efficiently. This is their function, because their ability to do these things is at the very least partly constitutive of what they are.

Here, we should keep the same points in mind as were made at the start of this section: When we judge an object to be valuable for someone, we need an explanation of what attitudes are fitting to direct toward it. Merely saying what properties make it valuable or what the precise bearer of the value ends up being, does us little good. It is also of little help to appeal to a reduction which seems to rob the concept of personal value of its connection to reasons for responses. In my view, this is what happens when we analyze the concept in terms of wellbeing and take wellbeing to be a purely nonnormative notion. Of course, there may be other nonnormative analyses that in fact do manage to carve out place for the concept of personal value in the broader normative domain. This is not something that we can or should conclusively rule out. Nevertheless, it would not be wrong to adopt a skeptical stance at this stage and remind ourselves that the burden of proof is wholly on proponents of descriptivism. Until they have assumed their theoretical burden, we may return our attention to the pattern of analysis that we have been using so far, which, incidentally, could be made compatible with most forms of descriptivism—should push come to shove.

Thomson (2008) and Foot (1985) seem to ground ideas about what is valuable for someone on ideas about what makes objects valuable as particular kind of things. Rosalind Hursthouse (1999), on the other hand, is among the philosophers who puts emphasis on wellbeing and understands this in nonnormative terms. Rosati (1996, 2008, 2009) is one of the more enthusiastic philosophers to invoke wellbeing in her discussions of personal value, but it is unclear whether she is equally reductive in her views as some of the former philosophers.

There is another way of making some of these points that invokes the notion of a thick value concept. A thick value concept is one that is either entirely or partly nonnormative. There is controversy as to whether there are any examples of the former, but examples of the latter are generally accepted in second-order value theory. For example, when we judge someone to be brave, we seem to be judging not only that this person is good, but that she is good in virtue of having certain nonnormative properties. The person is good, say, because she can overcome her fears in the pursuit of certain ends. Thin value concepts are those that, on the surface at least, appear to lack any nonnormative contents in this way. Examples include the concepts of being good and bad, which may be used without commitment to any story about why objects are good and bad. It seems fair to say that value in a way has often been assumed to be a family of thick value concepts. If the arguments soon to be presented are right, then this is a mistake. There is a reasonable understanding of value in a way that takes it to be a family of thin value concepts, just like the concepts of being good and bad.
The Fitting-Attitudes Analysis of Value in a Way

To establish value in a way as a distinct type of value, while staying clear of the explicitly nonnormative analyses that were mentioned in the previous section, we need to say something more about what the associated responses are meant to be like. The attitudes that are warranted by the presence of value in a way need to be distinguished from the attitudes that are called for by other types of value, including that of final value and that of instrumental value. To this end, we will soon move on to analyze the concepts of kind-value and personal value. We will secure their placement in the broader normative domain by making use of the fitting-attitudes analysis, which, again, aims to understand value in terms of the presence of reasons for responses. In so doing, we will come across several theoretical problems that need to be dealt with. What is clear is that regardless of exactly how the problems are solved, every single analysis we will be considering leaves the door open for the possibility that objects have final or instrumental value in a way. Before we start down this road we should briefly consider some substantive examples that illustrate the concepts in question. This might increase our confidence that there is something here to be analyzed in the first place.

There is an Aristotelian tradition that attempts to account for what makes people good by reference to what makes them good as human beings. Whether such an account holds any water is a matter of controversy that should not concern us now. The important thing is that if people can be good as human beings, then, in addition to whatever other values they may have, it seems reasonable to insist that they must also be good in a final way. By contrast, toasters, in so far as they are valuable as toasters, would typically be good only in an instrumental way. The idea, then, is that a human being and a toaster can be valuable as the kinds of things that they are, in a manner that also makes one fitting to favor for its own sake and the other fitting to favor for the sake of its effects. There are several ways of capturing this intuition, some of which are deliberately deflationist. For example, it may be pointed out that objects can always have two distinct types of value at once. People can therefore be good as human beings, while they also happen to be good for their own sakes. Objects can be good as toasters, while they also happen to be good for the sake of their effects. This observation is right, but it does not capture the relevant intuition.

The intuition is that there is some central difference between the value that accrues to people when they are good human beings and that which accrues to toasters when they are good toasters. What is more, the difference between the two values has not just to do with the nonnormative differences between human beings and toasters. Another way of capturing the difference would be to
understand it in terms of the varying weights that the respective values have. In other words, the value that accrues to people who are good as human beings may just be greater than the value that accrues to good toasters. This is also true as far as it goes, but it is important that we do not misunderstand what this means. Suppose we have two things before us, one of which manages to fulfill all the standards that are associated with good human beings and the other that manages to fulfill all the standards that are associated with good toasters. None of the objects is better than the other in terms of being the kinds of things that they are. Nevertheless, it seems reasonable that the value accruing to the human being has a different normative flavor from than the value that accrues to the toaster. One way of capturing this intuition is to suppose that the values fall on opposite sides of the distinction between the final and the instrumental.

The points that have just been made can be carried over to the case of personal value as well, which has all but fully recovered from the onslaught of Moore (1993/1903, p. 150) and is now enjoying increased attention in value theory. It seems reasonable to suggest that close friendships are good for anyone with social needs, since close friendships would enable them to have enriching experiences the absence of which would otherwise be missed. A pair of prescription glasses are also good for people with bad and correctible eye sight, at least partly for the same reason. Prescription glasses enable people to take part in activities that enrich their lives and add to their wellbeing. Yet, there is arguably something very different between these personal values. One interpretation has it that objects can have a personal value while also being valuable in a final or instrumental way. It might be, then, that in addition to being good for someone, close friendships also happen to be good for their own sakes. The same line of reasoning applies to prescription glasses, which might be good for someone while also being good for the sake of their effects. Values are friendly things that can share the same bearers even when the values belong to distinct types. Of course, this assumes that pluralism is correct and that it makes sense for things to be good or bad merely for their own sakes to begin with.

Another interpretation, also borrowed from the preceding paragraphs, suggests that the difference between the two objects should be understood by reference to their varying weights. It might be, in other words, that close friendships end up being better for someone than prescription glasses are. Again, this might very well be true, but there is a way of spelling this out that seems more faithful to the underlying intuition here. The idea is that there is a type of value which joins together the conceptual elements of personal value with those of final or instrumental value. This means that the values involved in the examples could also find themselves on opposite sides of the distinction between
the final and the instrumental. When close friendships are good for someone, then they are typically good for their own sakes. When prescription glasses are good for someone, on the other hand, then they are just good in an instrumental sense. Close friendships are special and something that we would not want to be without even in ideal circumstances. The same cannot be said for prescription glasses. First-order examples like these are never conclusive, but hopefully they are convincing enough motivate some of the analytic endeavors to follow.\textsuperscript{84}

Toward a Conditional Approach

Close attention will here be paid to accounts of both kind-value and personal value that are set in terms of fitting attitudes. The structure of these analyses will vary quite a lot depending on the theoretical choices we make along the way. The choices will be motivated by attempts to avoid a friction that often arises between the analyses and our common-sense view of kind-value or the specific ways in which the fitting-attitudes analysis is typically applied. For now, it suffices to recall that when we analyze value in terms of fitting attitudes, we want the bearers of the value to be the things toward which the relevant attitudes would be fitting. It should also be kept in mind that most of the applications of the fitting-attitudes analysis that will be presented throughout the upcoming subsections rely upon the same intuition about what is needed to understand both kind-value and personal value. Very roughly, the intuition says that when an object has value as a kind of thing, or has value for someone or other, this has implications for how the object should be treated in a certain kind of choice situation. The attempt to cash this out will occasionally lead us down some blind alleys, but with time it will also highlight some promising possibilities. We shall begin by looking at the following rival analyses:

\textsuperscript{84} Even though the interest of this chapter is not primarily a second-order one, the conclusions just reached should also have implications for first-order discussions within moral philosophy as well. It was mentioned earlier that there are philosophers, like Kraut (2011), that have taken inspiration from Geach and his followers while softening some of their arguments. Kraut maintains that the concept of final value might be meaningful, but it is also redundant when it comes to ethics and discussions about what we have most reasons to do. When we look to specific examples what matters most is invariably that objects are good or bad for us. The examples just offered are relevant to this claim, in so far as they show that value in a way might have a final character. It is therefore a mistake to treat the concept of final value as if it were in competition with that of value in a way for the center stage in ethics.
F9:  x is good in a way =_{df} if CHO, then it is fitting to favor\textsuperscript{B} x.\textsuperscript{85}

F10:  x is good in a way =_{df} it is fitting to favor\textsuperscript{A} that if CHO, then we favor\textsuperscript{B} x.

Favoring\textsuperscript{A} is hereafter understood as a positive response involving attitudes, while favoring\textsuperscript{B} is understood as a positive response involving actions. Whenever attitudes are mentioned in the discussions without the superscript, we should continue to understand them generically, as involving either. The intuitive reasons for involving distinct kinds of attitudes in these analyses are clarified by looking at the precise nature of the condition specified by CHO. In the case of kind-value, CHO should perhaps specify the situation of our choosing (recommending) a good member of a kind. It is therefore tempting to understand favoring\textsuperscript{B} in similar terms, as the choice (recommendation) of x. The issue of what sort of attitude is involved in favoring\textsuperscript{A} something is unfortunately less clear, but it seems rather natural to understand it in preferential terms. F9 can therefore be read as saying that x is good as a kind of thing if, and only if, we find ourselves in a situation of choosing (recommending) a good member of that kind, then it is fitting to choose (recommend) x. The rival analysis expressed by F10 entails that x is good as a kind of thing if, and only if, it is fitting to prefer that if we find ourselves in a situation of choosing (recommending) a good member of that kind, then we choose (recommend) x.

When we apply these patterns of analysis to personal value, it seems reasonable to understand CHO in more specific terms. Perhaps it should be taken to specify the situation of our choosing (recommending) a good member of a kind, where this kind is understood in terms of a class of objects that affect the health or wellbeing of a specific person. F9 can therefore be read as saying that x is a good for a person if, and only if, we find ourselves in a situation of choosing (recommending) an object that is conducive to the health or wellbeing of that person, then it is fitting to choose (recommend) x. The analysis expressed by F10 entails that x is good for a person if, and only if, it is fitting to prefer that if we find ourselves in a situation of choosing (recommending) an object that is

\textsuperscript{85} For the sake of simplicity and readability, the circularities which I take to characterize plausible applications of the fitting-attitudes analysis are left implicit in all the analyses to follow. It should be remembered that these circularities are different: On the one hand, the attitudes referred to in all those analyses are to be understood as implicitly representing objects as being good or bad for their own sakes—or as having properties that make them good or bad for their own sakes; On the other hand, the fittingness these attitudes is meant to be grounded in facts that actually afford the objects with their value.
conducive to the health or wellbeing of that person, then we choose (recommend) x. If this way of understanding CHO is on the right lines, then personal value becomes very closely related to kind-value. In fact, personal value will seem to be a specific type of kind-value, where the nature of the relevant choice situation is given a more detailed description. This is in accordance with what we imagined that proponents of descriptivism about these value concepts might want to say.

The narrow-scope analysis offered by F9 places the fitting favoring that is meant to ground the relevant value within the consequent of a material conditional. The result might seem to be that x cannot be said to have value in a way generally. Only when we are in the choice situation specified by CHO can x be good for someone or other or good as a kind of thing. This would indeed be strange, but the objection is based on a misunderstanding. The analysis offered by F9 reduces the relevant value to the material conditional and so an object can have value in a way regardless of whether we find ourselves in CHO. Of course, this might also seem to entail that F9 cannot represent a plausible application of the fitting-attitudes analysis. The reason is that the analysis in question makes it far too easy for an object x to possess the relevant type of value. After all, the conditional analysis expressed by F9 is trivially satisfied when the antecedent does not hold. This problem could only be avoided by giving F9 an indicative reading, so that it made claims about what would be fitting if CHO were the case. However, an aspect of the original problem survives, since F9 entails that x can have value in a way without this necessarily having any implications for what attitudes we should have. All it states is that as long as CHO is the case, then it is fitting to favor x.86

The wide-scope analysis offered by F10 avoids the problems just mentioned in a very different way, because it places the fitting favoringA outside of the material conditional. With regard to the analysis offered by F10, the problem is rather that the fitting favoring appears irrelevant when it comes to the specific value of the object x. The reason is that the favoring is apparently not fitted to x specifically, but rather to a conditional state of affairs that includes x. Philosophers who are well acquainted with the traditional applications of the fitting-attitudes analysis should recognize this immediately. For when the fitting-attitudes analysis is applied to account for the value of different objects, the attitudes that figure in it are ordinarily taken to be fitting to hold toward those objects specifically. In the case of kind-value, the analysis offered by F10

---

86 Gratitude is owed to Jens Johansson and Włodek Rabinowicz for discussing the implications of F9 with me and for pressing the objections just mentioned (in private communication).
entails that to judge x to be a good toaster is to ascribe value to the state of affairs that if we are choosing (recommending) a toaster, then we choose (recommend) x. Value in a way, which is typically assigned to objects rather than states of affairs, becomes a signatory value in the sense that it is only ascribed to hint at presence of some other good.  

The main upshot of what has been said so far, then, is that the structures of F9 and F10 force us to make revisions, either to our understanding of the fitting-attitudes analysis and its merits or to our common-sense view of value in a way. Whereas F9 implies that kind-value and personal value are values that are very easy to come by, F10 implies that the most relevant value in these contexts is ascribed, not to objects, but rather to the conditional states of affairs that include them. An additional worry meriting consideration in this context is that although both analyses could make the concept of value in a way susceptible to the distinction between final and instrumental value, F10 arguably does so in a rather unattractive and strange way. Notice, first, that there seems to be a meaningful question about the precise nature of the attitudes that figure in the consequent of each analysis. We know these attitudes are directed toward x, of course, but what else can be said about their respective characters? Here are some variations on F9 and F10 that attempt to answer the question:

\[ F11: \ x \text{ is finally good in a way } =_{df} \text{ if } CHO, \text{ then it is fitting to favor }^B x \text{ for its own sake.} \]

\[ F12: \ x \text{ is finally good in a way } =_{df} \text{ it is fitting to favor }^A \text{ that if } CHO, \text{ then we favor }^B x \text{ for its own sake.} \]

\[ F13: \ x \text{ is instrumentally good in a way } =_{df} \text{ if } CHO, \text{ then it is fitting to favor }^B x \text{ for the sake of its effects.} \]

---

87 Among the first philosophers to endorse the fitting-attitudes analysis for value in a way is Rawls, though he does so quite apart from any discussions of Geach and the challenge posed to the concept of final value. It is interesting to note that the application that he favors seems to suffer from the same misdirection problem that was mentioned moments ago. On this application, A is a good x if, and only if, “A has the properties (to a higher degree than the average or standard X) which it is rational to want in an X, given what X’s are used for, or expected to do, and the like (whichever rider is appropriate)” (1971, p. 399). This somewhat untidy analysis does not appear to capture the value of A as much as it captures the value of the properties of A (or perhaps the fact that A has those properties). Be that as it may, it seems as if the want to which Rawls refers in this analysis could also be understood as the kind of want that is involved in favoring objects either for their own sakes or for the sake of their effects.
**F14:** x is instrumentally good in a way $=_{df}$ it is fitting to favor$^A$ that if CHO, then we favor$^B$ x for the sake of its effects.

F11 offers a narrow-scope analysis and F12 offers a wide-scope analysis, they both involve the kind of favoring$^B$ that tends to be fitting to direct toward objects that are good for their own sakes. Conversely, while F13 offers a narrow-scope analysis and F14 offers a wide-scope analysis, they both involve the kind of favoring$^B$ that tends to be fitting to direct toward objects that are good for the sake of their effects. Now, if it is true that a meaningful question can be asked about the nature of the attitudes$^B$ that occur in the consequent of each analysis, the same should be true of the more complex attitudes$^A$ that figure in F10 and its variations. To be more precise, the question is whether the latter could also be specified along the lines that have just been explored, as involving the kinds of response that are fitting to direct toward objects with final or instrumental value. A positive answer would seem to recommend the following variations:

**F15:** x is finally good in a way $=_{df}$ it is fitting to favor$^A$ for its own sake that if CHO, then we favor$^B$ x for its own sake.

**F16:** x is instrumentally good in a way $=_{df}$ it is fitting to favor$^A$ for the sake of its effects that if CHO, then we favor$^B$ x for the sake of its effects.

If we wish to remain consistent in this exercise in combinatorics we would also saddle ourselves with the following more peculiar variations on F10:

**F17:** x is finally instrumentally good in a way $=_{df}$ it is fitting to favor$^A$ for its own sake that if CHO, then we favor$^B$ x for the sake of its effects.

**F18:** x is an instrumentally finally good in a way $=_{df}$ it is fitting to favor$^A$ for the sake of its effects that if CHO, then we favor$^B$ x for its own sake.

This unattractive and strange profligacy of different kinds of value in a way is surprising and a little worrying. Some of the variations on F10 appear to me somewhat artificial, which could mean that they do not capture concepts that are particularly common to our everyday talk and practices. It would be preferable, at least, if there was a way of accounting for the concept of value in a
way that does not burden us with this sort of profligacy. Luckily, there is a suggestion along these lines to be found, where the relevant favoring is taken to be more basic than any of these analyses would suggest. Before we move on to explore the benefits and potential problems of this account, though, some counterparts to the analyses above should perhaps be pointed out. Instead of being couched in terms of material conditionals, these counterparts rely on the controversial notion of a *conditional favoring*. This makes some of the resulting analyses immune to the structural problems mentioned in the preceding paragraphs. Consider first the following rival analyses of value in a way:

\[ F19: \text{x is good in a way}=_{\text{df}} \text{it is fitting to favor}^B \text{x given that CHO.} \]

\[ F20: \text{x is good in a way}=_{\text{df}} \text{it is fitting to favor}^A \text{the favoring}^B \text{of x given that CHO.} \]

F19 and F20 may appear equivalent to the analyses offered by F9 and F10, but close inspection will reveal that they are not. That is to say, the structural features of F19 and F20 do not match up to the logic of material conditionals, because they instead rely on the notion of a conditional favoring. When we conditionally favor an object, we should be taken to actually favor it. It does not matter whether the condition is satisfied, i.e., if the choice situation specified by CHO happens to obtain. In this sense, conditional attitudes behave very much like conditional bets. For when a person makes a bet on condition that CHO, it is wrong to think that she is not really making a bet at all. We should instead understand her as having made a bet that ends up being void in the event that CHO does not obtain.\(^88\) To understand what it means for a favoring to be void, it would be helpful if we now went on to apply these patterns of analyses to the two specific concepts of value in away on which we have so far focused.

In the case of kind-value, F19 can be read as saying that x is a good member of a kind if, and only if, it is fitting to choose (recommend) x, given that we are in the situation of choosing (recommending) a good member of that kind. The rival analysis expressed by F20 entails that x is a good member of a kind if, and only if, it is fitting to prefer that we choose (recommend) x, given that we are in the situation of choosing (recommending) a good member of that kind. In the case of personal value, F19 can be read as saying that x is good for a person if,

\(^{88}\) Gratitude is owed to Włodek Rabinowicz for suggesting this analogy to me (in private communication) and for mentioning the possibility of the given-that analyses that are discussed throughout this section.
and only if, it is fitting to choose (recommend) x, given that we are in the situation of choosing (recommending) an object that is conducive to the health or wellbeing of that person. The rival analysis expressed by F20 entails that x is good for a person if, and only if, it is fitting to prefer that we choose (recommend) x, given that we are in the situation of choosing (recommending) an object that is conducive to the health or wellbeing of that person. Since the conditional favoring\(^B\) referred to by F19 is clearly meant to involve an action in each of these applications, it becomes rather natural to suggest that CHO constitutes some sort of felicity condition for the relevant choice (recommendation). It is not that the choice (recommendation) of x is not made (or that it would not be fitting) unless CHO obtains, but that in the absence of CHO the action in question would not end up being effective.\(^89\)

However, since the conditional favoring\(^A\) referred to by F20 is clearly meant to involve an attitude rather than an action, it is much harder to account for the nature of the conditional favoring. More specifically, the challenge is to explain what it means for an attitude to become void, since felicity conditions only apply to speech-acts. Here the previously mentioned analogy with conditional bets only gets us so far.\(^90\) What all this means is that F19 has a big advantage over all the analyses offered so far in this section. It invokes a conditional favoring which might be mysterious in some respects, but which is still less mysterious than those invoked by the rival analysis F20. It also avoids the risk of making context sensitivity an analytic feature of value in a way, just as it avoids

\(^89\) There appears to be no greater mystery here, at least, than in the case of conditional bets. It seems to make perfect sense that a bet has been made, but that the bet is made unsuccessful by not being made in the right sort of circumstances. For more, see Austin (1962).

\(^90\) One strategy would be to try to reduce the conditional favoring\(^A\) to one that is not conditional. This is achieved by invoking ordinary preferences holding between two complex states of affairs: x is a good member of a kind \(=_{df}\) it is fitting to prefer [CHO and favoring\(^B\) x] over [not CHO and favoring\(^B\) x]. However, this reading makes it obvious that, much like its closest counterparts toward the start of this section, this analysis is vulnerable to the misdirection objection. The value of what is meant to be a concrete object is here analyzed in terms of a preference being fitting to direct to something else, namely the complex state of affairs that we find ourselves in the situation of choosing (recommending) a good member of a kind and choose (recommend) x. What this means is that value in a way still becomes a sort of signatory value, but one of a somewhat different kind than we originally feared. In fact, given the manner of reducing conditional attitudes that has just been presented, this may well end up being the case for some of the other analyses just presented as well. As long as the conditional favoring is not reduced, there is no misdirection. The value of what is supposed to be a concrete object is analyzed in terms of a choice (recommendation) being fitting to direct at that object specifically, which is really the desired result. Gratitude is once again owed to Wlodek Rabinowicz for helping me get clearer about these issues (in private conversation) and for mentioning the possibility that conditional attitudes might be reducible.
misdirection—reducing the value of an object to the value of something like a state of affairs. Consider now the following variations on both F19 and F20:

\[ F21: \] \( x \) is finally good in a way \( =_{df} \) it is fitting to favor\(^B\) \( x \) for its own sake given that CHO.

\[ F22: \] \( x \) is finally good in a way \( =_{df} \) it is fitting to favor\(^A\) the favoring\(^B\) of \( x \) for its own sake given that CHO.

\[ F23: \] \( x \) is instrumentally good in a way \( =_{df} \) it is fitting to favor\(^B\) \( x \) for the sake of its effects given that CHO.

\[ F24: \] \( x \) is instrumentally good in a way \( =_{df} \) it is fitting to favor\(^A\) the favoring\(^B\) of \( x \) for the sake of its effects given that CHO.

While the first two analyses involve the kind of favoring that is fitting to direct toward objects that are good for their own sakes, the second two involve the kind of favoring that is fitting to direct at objects that are good for the sake of their effects. Like their closest counterparts in the previous section, some of these analyses can be further developed by additionally specifying the kind of favoring\(^A\) they are meant to involve, thus yielding the following variations on F20:

\[ F25: \] \( x \) is finally good in a way \( =_{df} \) it is fitting to favor\(^A\) for its own sake the favoring\(^B\) of \( x \) for its own sake given that CHO.

\[ F26: \] \( x \) is instrumentally good in a way \( =_{df} \) it is fitting to favor\(^A\) for the sake of its effects the favoring\(^B\) of \( x \) for the sake of its effects given that CHO.

Once again, if we continue in this way we would seem to be saddled with something like the following analyses:

\[ F27: \] \( x \) is instrumentally finally good in a way \( =_{df} \) it is fitting to favor\(^A\) for the sake of its effects the favoring\(^B\) of \( x \) for its own sake given that CHO.
F28:  x is finally instrumentally good in a way =df it is fitting to favor\textsuperscript{a} for its own sake the favoring\textsuperscript{b} of x for the sake of its effects given that CHO.

At this point a reminder may be in order that the strange profligacy we have come across here and the preceding paragraphs is not just a result of questionable pedantry on my part. It is a seemingly unavoidable implication of the kinds of accounts that we have been pursuing throughout this section, based on different applications of the fitting-attitudes analysis of value. Whenever we choose to understand value in terms of reasons for attitudes there will always be questions about the nature of those attitudes. If we have an interest in making the analyses illuminating, and in setting value in a way apart as a distinctive family of values, these questions are not only legitimate to ask but necessary to answer. We will be required to say something about the multifarious ways in which objects can possibly be the fitting object of attitudes. The analyses explored so far invite answers that seem to open the door to a surprising variety of different value concepts, not all of which seem very interesting or likely to be represented in our evaluative talk and practices. However, some of the concepts appear to be plausible candidates for the concepts of final value in a way, which, if I am right, also have some intuitive support behind them. In the next section, we shall consider some alternative approaches to the ones just explored, which eschew any reliance upon conditional attitudes. The goal of these alternative approaches is to treat the attitudes that are called for by objects with value in a way as special kind of attitudes, with a unique intentional structure relative to the attitudes that tend to be referred to in these kinds of discussions. I will mention some of the problems that are associated with coming to grips with that intentional structure.

**Toward a Nonconditional Approach**

The analyses that we have looked at so far all rely on the intuition that that concepts like kind-value and personal value cannot be understood without reference to a certain sort of choice situation. The intuition is that although the value of something like a toaster should not depend upon the choice (recommendation) of a toaster, its value as a toaster still has a conceptual connection to those sorts of activities. The same is true if the toaster is good for someone, which has a connection to choosing (recommending) things that are conducive to the health or wellbeing of a person. Here we shall consider whether
value in a way is amenable to analyses that do not rely on this intuition. An alternative approach would be to treat the attitudes that are called for by objects with value in a way in simpler terms than we have done so far. This entails taking the attitudes for granted rather than understanding them in the conditional terms used earlier. We can then focus our conceptual work on clarifying and making distinctions in the way that properties are represented in the intentional contents of said attitudes. This is precisely how we tend to treat the responses that are called for by objects with final or instrumental value and so, if we are not tempted toward the conditional approach, this will seem like a more fitting strategy. Consider the following:

\[ F29: \quad \text{x is a good K} =_{df} \text{it is fitting to favor x as a K.} \]

\[ F30: \quad \text{x has value for P} =_{df} \text{it is fitting to favor x for the sake of P.}^{91} \]

What makes F29 and F30 special is that the attitudes that they refer to have a rather unique character relative to the attitudes that tend to figure in these sorts of contexts.\(^{92}\) It might be thought that the same was true of the attitudes considered throughout the last section, but this turned out to be largely illusory. Aside from the kinds of objects that those attitudes were fitting to direct at, their characteristics were already quite familiar to us. They were just the kinds of attitudes that it is fitting to direct toward objects that are good for their own sakes or for the sake of their effects. The attitudes referred to by F29 and F30 are

---

91 There may be a temptation to analyze personal value in terms of agent-relative reasons for responses. There are several interpretations of what it means for a reason to be agent-relative. All reasons are of course agent-relative in the trivial sense that they are reasons for someone or other. The more interesting notion is that of a reason that is agent-relative in that it is only a reason for a specific person to respond in certain ways. The idea, then, is that an object x has value for a person if, and only if, there are reasons for that person specifically to respond to the object in certain ways. This approach does not seem to me very plausible. To distinguish the types of value that there are, we need to look to the responses that they call for rather than the reasons that they are associated with. To focus too much on reasons would be much like trying to distinguish the types of value that there are by looking at the structure of their respective resultance bases. This seems like a red herring, especially from the perspective of a fitting-attitudes analysis of value and how this has traditionally been applied. For further discussions about the possibilities of analyzing personal value in terms of agent-relative reasons, see e.g. Rønnow-Rasmussen (2011, pp. 126–151) and Orsi (2015, pp. 73–79). For further discussions about the concept of agent-relative reasons as such, see e.g. Nagel (1970), Parfit (1984), Rønnow-Rasmussen (2009, 2012), and Ridge (2017).

92 Analyses of kind-value that look very much like this have been offered by other philosophers, among them being Skorupski (2010, pp. 82–87). See also Rowland (2016b, p. 215) for a different approach.
very different in this regard. In effect, the analyses in question place the attitudes that are called for by objects with value in a way on equal footing with any other type of attitude that has been involved in the fitting-attitudes analysis.

One question is whether the analyses can also accommodate the intuition defended in the previous subsections, that personal value can be understood as a special type of kind-value. This does not appear to be the case, unless there is some way of understanding the attitude appealed to in F30 as the kind of attitude that is appealed to in F29. For example, it might be argued that to favor an object for the sake of a person is the same as favoring an object as the kind of thing that would be conducive to the health or wellbeing of that person. However, this is not a very inviting prospect, and so accepting these kinds of analyses might force us to give up on the intuition that personal value should be understood in terms of kind-value. This does not preclude the possibility of connecting the two types of values on the first-order side of things, rather than by way of conceptual analysis. We might state that although the two types of values are conceptually distinct, it is nonetheless true that when an object is good for a person, then it is also good in virtue of being a good well-fare promoter relative to that person. The second and in the current context more pressing question is whether the analyses just put forward, given their relative simplicity, cannot be further developed to accommodate the concepts of final and instrumental value in a way. For the final time, consider now the following variations on F29 and F30:

\[ F31: \] x is a finally good K =\( \text{df} \) it is fitting to favor x for its own sake as a K.

\[ F32: \] x has final value for P =\( \text{df} \) it is fitting to favor x for its own sake for the sake of P.

\[ F33: \] x is an instrumentally good K =\( \text{df} \) it is fitting to favor x for the sake of its effects as a K.

\[ F34: \] x has instrumental value for P =\( \text{df} \) it is fitting to favor x for the sake of its effects for the sake of P.

While F31 and F32 involve the kind of favoring that is fitting to direct toward objects that are good for their own sakes, F33 and F34 involve the kind of
favoring that is fitting to direct at objects that are good for the sake of their effects. The main benefit of F29 and F30 and their variations is that they are not as vulnerable to the kind of structural difficulties discussed in previous sections. The main drawback is that to further clarify the analyses in question we need to trudge through some of the murkier waters in the philosophy of mind. The idea is of course that when we favor a human being for her own sake and when we favor an object as a toaster or even for someone, all these attitudes will differ in respect of their intentionality. They may not only be directed toward different objects, then, but the objects are meant to be represented in the three attitudes in subtly distinct ways. It is at this point that philosophers tend to become a bit tongue tied, for it seems almost impossible to describe clearly what those distinct ways might be. Rønnow-Rasmussen (2011, 2014) is among the philosophers to acknowledge the difficulty in coming to grips with what the attitudes appealed to in the sorts of analyses we have just been looking at. He starts out his discussions on the topic by offering an intuitive picture, saying that favoring something for the sake of someone is to favor it “with an eye toward this person” (2011, p. 56). The question is of course what this means. Rønnow-Rasmussen admits that “the prospects of specifying the characteristic features of this sort of attitude in a fully satisfactory way are bleak” (ibid, p. 56).

There are yet other complex forms of kind-value and personal value that are not discussed here. It might be possible for an object to be good for one person for yet another person. It is not too difficult to think of substantive examples that are meant to illustrate this possibility, even if grammar makes it difficult to express. Suppose that a relative of ours is having a birthday party. We have not seen this relative in many years and do not really feel like going. When hearing of our plans not to go, our parents are disappointed. They want us to go for the sake of our relative, because even though we do not know each other very well, they know that it would make our relative very happy. We eventually decide that we should go for the sake of our parents. However, we also happen to know that if we do not go to the birthday party for the sake of our relative, then our going will do no good. The relative will be able to tell that we are there mainly out of a sense of duty to our parents and thus, their wishes will go unfulfilled. There are many ways of cutting up the cake here, but it seems conceivable that we decide that we should favor going for our relative’s sake for our parent’s sake. The structure of this kind of attitude bears a resemblance to the structure of the other attitudes mentioned throughout this section. See Rønnow-Rasmussen (2011, pp. 57–60) for a further discussion of these issues.

Much of Rønnow-Rasmussen’s (2011, pp. 68–73) subsequent attempts at developing a richer story about the distinct roles that properties can play in the contents of our attitudes. In so doing he goes into depth about the concept of justifiers and identifiers. A justifier is a property that figures in the intentional content of our attitudes as the very thing that justifies our having them. Thus understood, justifiers have played an important part in the discussion above, as well as in the characterization of the fitting-attitudes analysis on which we here rely. An identifier, on the other hand, is a property that figures in the intentional content of our attitudes as the thing that uniquely identifies or demarcates the object to which it is directed.
Of course, if we accept the kinds of conceptual circularities which were mentioned in the introduction, then there are ways in which we can elaborate on this point. We might suggest that an object is favored as a kind of thing if the object is represented in the intentional content of the attitude as being good as that kind of thing. This can then be applied to the more complex attitudes that have been highlighted as well. An object is favored for its own sake as a kind of thing if the object is represented in the intentional content of the attitude as being good for its own sake as that kind of thing. An object is favored for the sake of its effects as a kind of thing if the object is represented in the intentional content of the attitude as being good for the sake of its effects as that kind of thing. We might continue this way by saying that an object is favored for someone if the object is represented in the intentional content of the attitude as being good for that someone. An object would then be favored for its own sake for someone if the object is represented in the intentional content of the attitude as being good for its own sake for someone. An object is favored for the sake of its effects for someone if the object is represented in the intentional content of the attitude as being good for the sake of its effects for someone.

The most reasonable analyses considered in the preceding section relied on a notion of conditional favoring that is controversial and so far, perhaps not very well understood. By contrast, the analyses offered here rely on a series of slippery distinctions describing the way that things should figure in the intentional contents of our attitudes. Whether we ultimately prefer the former analyses, or the more basic approach explored here, is likely to hinge on our general interest in, and instincts about, conditional notions, rather than wider questions in the philosophy of mind. However, it should be noted that even if we were tempted to adopt one of the analyses from the preceding section, we would still need the slippery distinctions in question. We will still have to trudge through the murky waters of the philosophy of mind if we are make sense of the kinds of attitudes that it is fitting to direct toward objects with final or instrumental value. We saw earlier that even those conditional attitudes that might be called for with objects with value in a way can have this complex sort of character. What this means is that although the basic approach is saddled with some difficult problems, they are also very general problems. In fact, they seem to throw doubt over any attempts to capture the varieties of value by relying on a corresponding set of properly discerning attitudes.

Finally, I wish to say something about the extent to which I have relied upon the fitting-attitudes analysis of value. The discussion started out with some substantive examples that were meant to illustrate the possibility that objects can have final or instrumental value in a way. The substantive examples were then
given additional support by observing how the values in question are mirrored by attitudes that we might adopt. In the face of the substantive examples, it seems reasonable to suggest that some of those possible attitudes will even be fitting. This establishes a burden of proof against potential skeptics. For they would have to argue that it is impossible for objects to have final or instrumental value in a way, irrespective of how it might be possible or even fitting to favor the objects. There would be something strange about this argument. My view is of course that there are no strong reasons to reject the fitting-attitudes analysis in its circular form. However, it seems clear that even if we did have such reasons, the burden of proof against potential skeptics would still stand. It would just not be quite as insurmountable as we may have hoped.\textsuperscript{95} One way in which potential skeptics might get around that burden of proof is by performing two theoretical tasks: First, they could offer up an alternative analysis of value in a way that respects and accommodates its inherent normativity to the same degree as the fitting-attitudes analysis; second, they could clarify how and why their alternative analysis of value in a way precludes the possibility that it can also come in either a final or an instrumental form.

**Concluding Remarks**

This chapter has focused on an influential objection to the concept of final value, which takes issue with the idea that objects can be merely good or bad for their own sakes. The objection states that this idea is at best an invention of philosophers, since objects could only be good or bad in a way. For the sake of simplicity this was taken to mean that objects can only have value as kinds of things or that they can have value for someone. For all its influence, the arguments behind this objection turn out to be rather weak. The first argument relies upon a proposed method for distinguishing between attributive and

\textsuperscript{95} Suppose we deny that there is any relation between value and reasons for responses. It is not just that value cannot be understood partly in terms of the presence of reasons, then, but not even a necessary bi-conditional could be established between the two. In one way, this would make it more difficult to argue in favor of the concepts of being finally and instrumentally valuable in a way, but in another way, it might make it easier. Absent any analysis of value concepts, all we can do is rely on the brute intuition and start table-thumping. It seems entirely conceivable to me, or at the very least not obviously confused, to judge objects to be finally or instrumentally valuable in a way. The question is what the supposed objection is that is meant to convince me that this is a mistake.
predicative uses of value concepts. As opposed to predicative uses, attributive uses of value concepts always relate values to kinds and perhaps the standards that those kinds are essentially associated with. The method proceeds by observing how value concepts tend to behave in arguments and inferences. Geach thinks that there are certain patterns of validity that make it obvious how closely related value is to kinds and their standards. However, we saw that there is more than one way of explaining the patterns of validity that he has in mind. The other explanation, which is more conciliatory to our intuitions and everyday talk and practices, suggests that value concepts involve ambiguities.

Some uses of value concepts are indeed essentially attributive, just as Geach proposes. Such uses cannot be understood without reference to some associated kind, although, as I have mentioned, there may be legitimate doubts about how determinate the standards that correspond to those kinds must be. Other uses of value concepts, however, are essentially predicative and therefore do not have to relate to any kinds in this way. Precisely which use is relevant to a specific case, i.e. whether a given value judgment involves a value concept of an essentially attributive or essentially predicative character, is made clear by linguistic and other contextual factors. That there may also be very opaque cases in this regard need not be especially problematic, in so far as both attributive and predicative uses of value concepts share a common core. Here the second argument in favor of the objection to final value gains relevance. Geach seems to worry that if there is an ambiguity when it comes to value concepts, then it might turn out that their different uses have little or no connecting tissue. The terms that are used to denote them would seem to be mere homonyms. To respond to the second argument supporting Geach’s objection to final value, the fitting-attitudes analysis was applied to value in a way. It was shown throughout the second part of the chapter that there are ways to account for this family of concepts in normative terms as well. The connecting tissue that bridges final value and value in a way is made from the presence of reasons.

What is more, all the applications of the fitting-attitudes analysis that have been brought up and discussed seem to leave the door wide open for the possibility that value in a way should come in both the final and the instrumental variety. This means that the concept of final value is not only one of the most prominent to figure in value theory, it is arguably one of the most fundamental concepts to be represented in our evaluative talk and practices. If this is right, then we can respond to the arguments offered by Geach and his followers in the following manner: The objection that things cannot be merely good or bad for their own sakes is not correct and even if it were, it would not actually matter. This response, if sound, should inspire a more conciliatory turn.
in value theory. For it offers us an assurance that philosophers who are primarily interested in exploring what it means for objects to have final value need not preface their work with lengthy defenses of this concept. In other words, even if it is the case that all objects with value end up being good or bad in a way, whatever has been established by philosophers about the concept of final value could still be relevant to discussions about certain types of value in a way.
When we start to think carefully about final value and its relations, certain observations are immediately invited. Among them is the observation mentioned in the introductory chapter, that final value, like all kinds of value, depends upon other things. This view is so commonly held that it could be treated as part of the shared ground on which nearly all discussions in value theory take place. The crucial question is of course how it should be understood in detail: What does it mean to say that final value depends upon other things in the world? In what sense is an object’s being good or bad for its own sake to be explained by the presence of something else? We have already distinguished between two different relations that vary in their scope: The dependence of final value was taken to have the widest scope, in that it involves all the factors that are relevant to the explanation why certain objects are good or bad for their own sakes. This includes the makers of final value and other facts that play a less direct role. Evaluative resultance was understood to have a narrower scope, picking out the relation between final value and the factors in virtue of which, or as a result of which, objects end up being good or bad for their own sakes. In other words, evaluative resultance only involves the makers of final value.

Although evaluative resultance is a narrower relation than the dependence of final value, it should be noted that it is also subject to more fine-grained distinctions. When philosophers talk in terms of the factors that make objects good or bad for their own sakes, they tend to gloss over the distinct roles that the relevant kind of factors are meant to be playing in narrow normative explanations. The makers of final value are not only directly responsible for whether objects have value, they are also directly responsible for determining
whether objects are good or bad (i.e. for the polarity of the resulting value), in what way objects are good or bad (i.e. for the type of value that results), to what degree objects are good or bad (i.e. for how much weight the resulting values have), what specific objects are good or bad (i.e. for where the resulting values are ultimately located). It is entirely possible, and it is worth keeping in mind, that each of these roles could be played by distinct factors in narrow normative explanations. What this means is that although evaluative resultance is indeed a more fine-grained relation than the dependence of final value, it, too, functions as a bit of an umbrella term, which gathers different relations and mechanisms.

We shall ignore such complications at the outset and try to make do only with the distinction between the dependence of final value and evaluative resultance. This distinction is perhaps a bit rough, but it is sufficient given our present purposes. The following chapter aims to explore these relations a bit further and to look closely at the worries that may be invited by their explication. In the introductory chapter to this work, the relations in question were briefly highlighted by way of examples. We imagined walking through a museum and overhearing a friend say of two numerically distinct objects, such as a couple of paintings, that one of them is good for its own sake and that the other is bad for its own sake. We proceed to ask our friend to explain why she thinks that one painting is good for its own sake and the other bad for its own sake. She explains that nothing makes them so. The paintings are identical in every relevant respect, she says, save for the fact that they differ in terms of their final value. It was suggested that this is a very strange answer and that the value judgments expressed by our friend run contrary to at least two deeply ingrained intuitions that many of us share about final value and its dependence upon other things. It would be helpful if we reminded ourselves, in rough terms, what those intuitions were and gave them some appropriate labels:

*Priority: An object cannot be good or bad for its own sake, unless the object has other properties that explain why the object has final value.*

---

96 This might seem to be in friction with the possibility that properties themselves can be bearers of final value, but I do not think this is right. In so far as properties are good or bad for their own sakes, this must be in virtue of the properties that those properties have. If we are uncomfortable with this way of speaking, we can instead say that insofar as properties are good or bad for their own sakes, this must be in virtue of some aspect of their nature.
**Similarity:** Two objects that are exactly similar in terms of their nonnormative properties must be exactly similar in terms of their final value.\(^97\)

The question was left unanswered whether there is some analysis of the dependence of final value that can capture each of these intuitions. The first part of the present chapter aims to explore this possibility more closely. Several proposed analyses will be rejected for either for being too strong or too weak to adequately capture the intuitions in question. The argument is made that evaluative supervenience, although useful and interesting for its structure and clarity, is not able to adequately capture the thick notion of priority or explanation that underlies the intuitions just listed. The main problem is that the dependence of final value seems to display features like irreflexivity, asymmetry, and hyperintensionality, which evaluative supervenience lacks. The later sections of the chapter is dedicated to some worries that are associated with the analyses of the dependence of final value that are adequate in this sense. One such worry is that the dependence of final value is strong enough to invite a reductionism about final value, whereby this value is identified with a nonnormative property. The argument relies upon an extensional principle saying that if two predicates end up being necessarily coextensive, then they must pick out the same property. Doubts will be raised about the principle, although, admittedly, much of the discussion lands in a stalemate.

Another worry is that the thick notion of explanation or priority involved in the dependence of final value lies very close, and may in fact be equivalent to, the metaphysical concept of *grounding*, which some philosophers argue is either incoherent or not very useful from a theoretical point of view. Here the argument is made that neither of these objections is very worrying. Even if the dependence of final value was revealed to be incoherent, moreover, this does not mean that an account which relies upon the notion is necessarily incorrect. In fact, it may just be that our evaluative talk and practices, as they currently look, involve notions that are incoherent. The argument is also made, then, that the thick notion of explanation or priority is not useless regarding the project of attaining a better understanding of our evaluative talk and practices. In connection with this discussion, some very general remarks will be made about conceptual issues and their connection to metaphysics that help clarify the

\(^{97}\) We will later see that this intuition should perhaps be spelled out in terms of a similarity of relevant features. We will also return to the issue of how the concept of relevancy should be understood.
project which is pursued in this work. Finally, I wish to note that the three challenges concerning the dependence of final value that have just been described are closely related. The point I wish to make is that if final value depends upon nonnormative properties, but this dependence cannot be fully reduced to other relations, such as evaluative supervenience or identity, then both final value and its dependence may end up looking rather mysterious. It is this mystery with which the chapter is primarily concerned. The aim of the chapter is not to dissolve the mystery as much as to highlight it, and address how much of a problem it constitutes for any accounts that invoke notions like dependence and evaluative resultance.

Analyzing Evaluative Dependence

The dependence of final value is occasionally analyzed in terms of evaluative supervenience. In this section, it will be shown that evaluative supervenience, at least on certain understandings of it, is not able to fully capture the priority intuition or the similarity intuition. What is involved in the dependence of final value, as well as evaluative resultance, is thereby illuminated. Now, the general discussions that have taken place over the past decades about how to understand supervenience relations are terrific examples of how philosophy progresses. We have a better idea than ever before of what options there are and of the benefits and disadvantages that they are associated with (e.g. Kim, 1993; McLaughlin & Bennett, 2014). An entire book could easily be written on the topic, with each chapter dedicated to carefully highlighting the benefits and disadvantages that are associated with a specific pattern of analysis. Unfortunately, this is not that book. Many options must be left unexplored and those that are mentioned will not be treated exhaustively, with all their limits and implications brought to light. Instead, a selection among the patterns of analysis must be made, based on their assumed relevance to the dependence of final value, as well as their perceived faithfulness to the everyday descriptions of the intuitions just

98 Concepts of supervenience have, in a more general sense, been employed across manifold philosophical areas to model the relations between many kinds of properties, events, and laws. For this reason, the concepts are now seen as a standard tool for philosophers within the analytic tradition. For a brief historical overview of the concept, with references to the most important sources, see Kim (1993) and MacLaughlin & Bennett (2014, section 2).
We shall start out by looking at a wide conception of evaluative supervenience, expressed by the analysis below:

**SI:** Final value depends upon nonnormative properties \(=_{df}\) for all possible worlds \(w_1\) and \(w_2\):

1. If \(w_1\) contains a distribution of final values, then \(w_1\) must also contain a distribution of nonnormative properties, and
2. If \(w_2\) is identical to \(w_1\) in terms of its distribution of nonnormative properties, then \(w_2\) must also be identical to \(w_1\) in terms of its distribution of final value.

There might be an intuition here that SI cannot help but be too wide in its scope. This is because of its global character—its focus on the behavior and characteristics of entire possible worlds. It seems like a natural enough thought that when we ordinarily explain why certain objects end up being good or bad for their own sakes, we will not always have entire possible worlds, or anything remotely as grand, in mind. Even broad normative explanations will tend to concern themselves with objects of a more local character, such as a painting, a person, a pen, and so on. It seems just as natural to suppose that in so far as global supervenience analyses like SI are correct, this should be because a more

---

99 Although the concept of supervenience was arguably first highlighted in value theory, many types of supervenience have since been formulated against the background of a discussion about the nature of consciousness or natural laws, rather than anything that might be of direct relevance to the normative domain (e.g. McLaughlin & Bennett, 2014, section 4.3.).

100 Ever since philosophers like Horgan (1982), Currie (1984), Kim (1984), Petrie (1987) and others first put explicit emphasis and elaborated upon the concept of global supervenience as a theoretical alternative, it has typically been formulated exclusively in terms of trans-world similarities. In other words, global supervenience has rarely incorporated the first part just offered, which states that for a world to contain final value in the first place, it also needs to contain other things. It is thought to suffice that if two worlds, \(w_1\) and \(w_2\), are identical in terms of their distributions of nonnormative properties, then they must also be identical in terms of their distributions of final value. See McLaughlin & Bennett (2014, section 4.3) for a general overview of how this seemingly simple idea can be cashed out in even closer detail. The first part of the characterization is included here to make the global analysis fit better with the intuitions described in the introduction. This also means that the global analysis will be more structurally similar to the local understandings soon to be discussed, which, to my mind, facilitates the understanding of both patterns of analysis. We should also keep in mind that although many supervenience analyses tend to be formulated in terms of naturalistic properties, they are here formulated in terms of the broader category of nonnormative properties. That we should want to be this general in our approach is motivated, in part, by difficulties in coming to grips with what naturalistic properties are, as well as the suspicion that final value could depend upon properties of other kinds.
local counterpart to such analyses is also correct. A first stab at adjusting our understanding of the dependence of final value accordingly might yield something like the following weak supervenience analysis:

S2: Final value depends upon nonnormative properties \(=_{df}\) for any possible world \(w_1\) and any objects \(x\) and \(y\) in \(w_1\), i) if \(x\) has a final value in \(w_1\), then \(x\) must also have a nonnormative property in \(w_1\), and ii) if \(y\) in \(w_1\) is identical to \(x\) in \(w_1\) in terms of its nonnormative properties, then \(y\) is identical to \(x\) in terms of its final value.\(^{101}\)

The objects mentioned by S2 could be something almost as grand as entire possible worlds, but they need not be. They could also be objects of a far more local character, such as a painting, a person, a pen, and so on. The analyses expressed by S1 and S2 appear to be logically independent from one another (Kim, 1987, p. 319). Suppose there are two objects, \(x\) and \(y\), that are situated on opposite sides of the border between two possible worlds, \(w_1\) and \(w_2\). It is compatible with S2 that \(x\) and \(y\) should be identical in all relevant aspects, save for the fact that one of them is good for its own sake and one of them is bad for its own sake. It follows that \(w_1\) and \(w_2\) can be identical in terms of their respective distributions of nonnormative properties, without thereby being identical in terms of their distribution of final values. We therefore have a counterexample to the idea that there is an entailment from S2 to S1. Suppose next that \(w_1\) contains two objects, \(x\) and \(y\), that have the same nonnormative properties, but differ in respect of their final value. This would obviously violate the condition specified by S2, but as long as the same situation holds in every possible world in which two objects have the same nonnormative properties as \(x\) and \(y\) in \(w_1\), there is no violation of S1 (cf. Petrie, 1987).

Although it undoubtedly represents an improvement over S1, there is a worry here that S2 ends up saying too little. More specifically, S2 seems far too weak to adequately capture the general intuitions that seem to underlie the dependence of final value. One of the reasons why S2 is too weak is that it is too

\(^{101}\) Just like global supervenience, Kim (1984, 1987) is also among those most responsible for first drawing attention to the weak analysis in philosophy. Unlike global supervenience, which we have formulated very differently from the way he first formulated it, the weak formulation of supervenience that occurs here is essentially unchanged except in terms of language. The same is true for the strong formulation of supervenience which is soon to be discussed. They are equivalent to what he had in mind.
conservative, modally speaking. It manages to capture the fact that if an object has final value in this or any other possible world, then it must have other properties that explain why this is so. However, the intuition is also that if an object is identical to another object in this or another possible world in terms of its nonnormative properties, then they must share the same final value. In other words, the evaluative similarity that needs to hold between two objects that are identical in all other respects does not involve mere correlations that hold on particularly sunny days but are instead meant to offer us guarantees (e.g. Kim, 1984, p. 159; Strandberg, 2004, p. 241). The thought experiment offered in the previous paragraph illustrates quite clearly what is at stake, and in what sense S2 betrays the intuitions mentioned in the introduction (Kim, 1984, p. 160; cf. Blackburn, 1985, p. 134). It should be inconceivable that objects situated on opposite sides of a modal border can be identical in all aspects, save for the fact that one of them is good for its own sake and the other is bad for its own sake. A strong supervenience analysis would make sure to specify that the two intuitions that we are hoping to capture hold true out of necessity:

**S3:** Final value depends upon nonnormative properties $=_{df}$ for any possible worlds $w_1$ and any object $x$ in $w_1$, i) if $x$ has a final value in $w_1$, then $x$ must also have some nonnormative property in $w_1$, and ii) for any $w_2$ and any $y$ in $w_2$, if $y$ in $w_2$ is identical to $x$ in $w_1$ in terms of its nonnormative properties, then $y$ in $w_2$ is identical to $x$ in $w_1$ in terms of its final value.

That S3 entails S2 but not the other way around seems intuitively obvious, but the relation between S3 and S1 is more open to controversy. For when he first compares these along the lines of S3 and S1, Kim (1984, pp. 168) explicitly states that global supervenience analyses are logically equivalent to strong supervenience analyses. He (1987, pp. 317–318) later admits that this can be shown to be false by several plausible counterexamples.\(^{102}\) Among the first counterexamples offered up in the literature is the one by Petrie (1987). He asks us to suppose that the object $x$ has the nonnormative property N and the final value $V$ within $w_1$, while $y$ only has the nonnormative property N within $w_1$. On the other side of the modal border, within $w_2$, $x$ has the nonnormative property N but not the final value $V$, while $y$ does not have N. It is clear that, as they have just been described, these possible worlds fulfill the demands of S1, but not

---

\(^{102}\) Among those to correct him on this issue in the literature, Kim mentions Hellman (1985, pp. 607–616) and Bacon (1986, pp. 163–176).
of S3. As far as I am aware no similar counterexample can be formulated that fulfills the demands of S3 without fulfilling the demands of S1.

There are at least two big problems with S3: First, there is the general worry that any analysis of dependence may end up capturing something true about final value, only by saying too much. There are many factors that we would not want to include even in a broad explanation of final value that would nonetheless be deemed relevant by S3. In fact, any minuscule difference between objects, no matter how insignificant or seemingly irrelevant, would free us from the commitment of assigning identical distributions of final value to them. On a related note, there are many nonnormative properties that we would not regard as relevant to the case at hand and whose presence would trivialize the dependence of final value.\footnote{Gratitude is owed to Włodek Rabinowicz and Toni Rønnow-Rasmussen for pressing this objection (in private communication).} An example would be the property of being identical to x and the world being identical to $w_1$. For any y in any $w_2$ that has this nonnormative property it is trivially true that y (which is identical to x) has in that world (which is identical to $w_1$) the same final value as x in $w_1$. One possible solution is to formulate S3 in terms of naturalistic properties, rather than nonnormative properties. Naturalistic properties are roughly the kinds of causally effective properties that may be referred to within the sciences. My view is that we should be wary of this solution because we should admit to the possibility that final value may depend upon other properties as well.\footnote{A superstitious person might say of another person that he is bad because he is in league with the devil. The property of the person being such that he is in league with the devil is not a naturalistic one. We might object on first-order grounds to the explanation in question, but that is of course beside the point. It should also be mentioned that just because we here focus on the dependence of final value upon nonnormative properties, we do not therefore disregard the possibility that final value can come to depend upon other normative properties. Indeed, this is a topic which we return to at several points in the upcoming discussion. We only assume here that although a final value may depend upon normative properties, which in turn depend upon yet other normative properties, there must be at bottom some nonnormative property which supports the entire normative structure.}

Another way of getting around the difficulty would be to understand the dependence of final value in terms of a similarity of relevant nonnormative properties. Objects need not be exactly alike to ensure that they have the same final value, then, but rather identical in respect of a narrower class of properties.\footnote{A related point is that we may not want the dependence of final value to be treated as a too open relation. For the dependence of final value to be open in the sense I have in mind means that when a final value depends upon a subset of properties within a superset of nonnormative} The question then arises how on earth we should explicate the
relevancy condition here: What defines the narrower class of properties that should be of interest? What are the characteristic earmarks that indicate whether a factor is relevant to a broad normative explanation? We are here operating under the assumption that broad normative explanations of final value may involve factors of roughly two distinct types: The first way in which a factor can influence final value is by being among the things in virtue of which, or as a result of which, an object ends up being good or bad for its own sake. The second way in which a property can make a difference for final value is by being among the things that determine how other factors figure in the relevant resultance base to begin with. Given this, the most natural suggestion seems to be that a property is relevant if, and only if, it belongs to one of these two types of factor. There is still the worry that this does not constitute a plausible general condition, and that our intuitions about what is relevant are ultimately tied to a prior understanding of dependence and what it entails. My suspicion is that this is right and that an accurate analysis of dependence cannot avoid circularity.

A different problem is that if we take the analysis expressed by S3 at face value, then the dependence of final value upon nonnormative properties looks to be a transitive, reflexive, and antisymmetric affair (Kim, 1984, p. 166): Its being transitive means that if a final value x supervenes upon the nonnormative property y and y in turn supervenes upon the nonnormative property z, then x supervenes upon y; Its being reflexive means that every final value x supervenes upon itself; Its being antisymmetric means that it entails neither symmetry nor asymmetry: If x supervenes upon y and x and y are identical, then there is a guarantee that y supervenes upon x as well. If x supervenes upon x and x and y are distinct from one another, then there is no guarantee that y supervenes upon x. The question to be considered is whether these are formal properties that we would like to assign to the dependence of final value upon nonnormative properties as well. I do not think so, for the dependence of final value upon nonnormative properties seems to me to impose a partial order upon its domain.

Another amendment would be to say that the only relevant nonnormative properties are universal ones, i.e. the properties that do not involve relations to specific objects or their identities (i.e. we exclude the property of being identical with x, which was referenced above). We will thereby have involved in our idea of evaluative supervenience the separate idea of universality (Hare, 1984, p. 3). This is meant to be understood in second-order terms and not as the thick requirement to universality often found in first-order theory. When some philosophers say that value is universal, they are making the ethical point that equal value should be assigned to all living creatures or that equal value should be assigned to the preferences of all living creatures, or something similar. This is not what interests us here.
meaning that it should be a transitive, irreflexive, and asymmetric affair. To see this, we need only reconsider the thick notion of explanation and priority which is meant to be involved here.

When our friend judges that a painting is good for its own sake, we do not expect her to say that this is partly to be explained by its being good for its own sake. When she judges that the other painting is bad for its own sake, we do not expect her to say that this is to be partly explained by its being bad for its own sake. When she says that the painting is good for its own sake is to be explained by its having certain colors and shapes, we also do not expect her to say that the value likewise explains the presence of the colors and shapes. These are simply not the kind of stories we are interested in, even when we are in fact asking for a broad normative explanation involving all the factors that make a difference for final value. Instead, we would expect our friend to point toward some of the other properties of the painting, with an implicit understanding that those properties are in some sense metaphysically prior to the final value which is explained by them. The dependence of final value is in this sense a nonreflexive and asymmetric affair. For this to be reflected in our understanding, we need to amend the relations captured by S3 so that they are couched in terms of some nonreflexive and asymmetric notion. We need a robust analysis of the dependence of final value that goes beyond evaluative supervenience. The simplest way of doing this is to just add an explicit mention of explanation within the analysis itself:

107 A mere strong supervenience analysis of evaluative dependence would leave us even worse off than this, for it would also entail that final value must partly supervene upon any other necessary fact. That one painting is good for its own sake will always supervene, at least in part, on its having been made by Vermeer or not having been made by Vermeer. That happiness has final value is at least partly to be explained by the fact that 2+2=4. This is not the kind of normative explanation we are interested in, clearly.

108 Kim himself admits as much and writes that supervenience is not explanatory or metaphysical in a deep sense: “rather, it is a ‘surface’ relation that reports a pattern of property covariation, suggesting the presence of an interesting dependency relation that might explain it” (1993, p. 167). He also gives hints about what this dependency relation might be like. Kim suggests that when physicalists rely upon supervenience to characterize the dependence of mental states on physical states, “they pretty clearly have in mind an asymmetric relation: they would say that their thesis automatically excludes the converse dependence of the physical on the mental” (1990, p. 13). It is this interesting dependency relation that we see reflected in our evaluative talk and practices as well, and which we are now attempting to conceptually approach. To this end, evaluative supervenience gives us a nice formal structure from which to start, but it fails in respect of its contents. Curiously, it seems that some philosophers would see all this as a benefit of evaluative supervenience. Lewis might be among them, as at one point he commends supervenience relations generally for being “unencumbered by dubious denials of existence, claims of ontological priority, or claims of translatability” (Lewis 1999, p. 29).
S4: Final value depends upon nonnormative properties \( =_{df} \) for any possible worlds \( w_1 \) and any object \( x \) in \( w_1 \), i) if \( x \) has a final value in \( w_1 \), then \( x \) must also have some nonnormative property in \( w_1 \) such that \( x \) having this nonnormative property explains its final value, and ii) for any \( w_2 \) and any \( y \) in \( w_2 \), if \( y \) in \( w_2 \) is identical to \( x \) in \( w_1 \) in terms of its relevant nonnormative properties, then \( y \) in \( w_2 \) is identical to \( x \) in \( w_1 \) in terms of its final value.\(^{109}\)

S4 understands the dependence of final value upon nonnormative properties as imposing a partial order upon its domain. S4 also differs from the other analyses by involving a thick notion of explanation, which is hyperintensional in nature. This means that the sense of explanation that is relevant here cannot be captured by wholly intensional accounts specifying what must be the case for something else to be the case. Correlations, be they contingent or necessary, are not enough. Sets of entailments are not enough.\(^{110}\)

\(^{109}\) At first, Blackburn (1971) offers up an understanding of evaluative supervenience that explicitly precludes value from being identical to the properties on which it supervenes. In a revisiting paper, Blackburn (1984) then takes a step further and offers up an understanding that could be interpreted as including the notion of metaphysical explanation or priority that we have in mind here, but he speaks instead of the naturalistic underlying the normative. This seems to indicate that whenever Blackburn spoke of supervenience, he, at least, had in mind something significantly thicker than the formal relations highlighted by philosophers like Kim. See also Bricker (2006), who puts forward an analysis of metaphysical grounding in terms of fundamentality and supervenience, in a manner that is similar to the one just defended above.

\(^{110}\) Arguably, there could not be a weak version of the robust analysis, which incorporates a thick notion of explanation, but which does not offer a guarantee of evaluative similarity across possible worlds. Suppose a painting in this world is beautiful because it features certain colors and shapes. The beauty of the painting is explained, in a broad sense, by the presence of the colors and shapes that it has. Let us then imagine a painting in another possible world that displays the very same colors and shapes. The question to be asked is whether the second painting could fail to be beautiful. The idea has just been mentioned is that this possibility is blocked by the formal features of our evaluative talk and practices. When we look closely at our evaluative talk and practices, they reveal that the concept of being beautiful commits us to the relevant sort of evaluative similarity. However, there is also a feature of the concept of broad explanations that ensures us against this possibility. We are assuming that the beauty of the painting in this world is broadly explained by the painting having certain colors and shapes. If it turns out that there is a painting in another world with the very same colors and shapes that is not beautiful, then it turns out that the assumption is wrong: the beauty of the painting in this world is not broadly explained by the painting having certain colors and
structure of the previous analyses, then, the inclusion of the thick notion of explanation means that we have taken several steps beyond mere evaluative supervenience. I take the notion of explanation that is embedded in this analysis is connected to a notion of metaphysical priority. When a final value is judged to be explained by the presence of certain nonnormative properties, then those nonnormative properties are, ipso facto, judged to be prior to the final value that they give rise to. Consider finally the following global version of S4:

S5: Final value depends upon nonnormative properties =df for all possible worlds \( w_1 \) and \( w_2 \) i) if \( w_1 \) contains a distribution of final values, then \( w_1 \) also contain a distribution of nonnormative properties such that \( x \) having the distribution of nonnormative properties explains its having the distribution of final values, and ii) if \( w_2 \) is identical to \( w_1 \) in terms of its relevant distribution of nonnormative properties, then \( w_2 \) must also be identical to \( w_1 \) in terms of its distribution of final values.

S5 focuses on the behavior and characteristics of possible worlds but is also formulated in terms of the distribution of nonnormative properties that explain the presence of a distribution of final values. It seems obvious that S5 is logically equivalent to S4. After all, counterexamples along the lines of that which was brought forward by Petrie are no longer applicable here. As long as N is assumed to metaphysically explain V, we cannot conceive of \( x \) as having the nonnormative property N and the final value V within \( w_1 \), while \( y \) only has the nonnormative property N within \( w_1 \). After all, the conceivability of this kind of gap between N and V would entail that the former cannot metaphysically

shapes. Something else that is relevant to the broad explanation must have been left out of the picture, something that is particular to the one possible world and not the other.

Horgan (1993) suggests that when there is supervenience and we can give a deeper story as to why, then there is also superdupervenience. This question might be asked whether the robust analysis brings us closer to this ideal. After all, if the robust analysis is correct, then we know why strong supervenience also holds. Unfortunately, this is probably not quite enough. Horgan (1993, p. 566) further states that in order to achieve superdupervenience, we would also need a story that is materialistically respectable. Presumably, this does not just mean that the robust analysis ascribes metaphysical priority to the material or the physical. The demand is that aspects of the robust analysis itself are amenable to materialistic or physicalist explication. Giving a detailed story as to why the concept of final value has the formal features that it has might be partly materialistic or physicalist. However, it should be admitted that the formal features themselves, in so far as they involve commitments to metaphysical explanation and priority, might not be so amenable (cf. Fine, 1994; Rosen, 2010).
explain the latter in a broad sense. This kind of case is a nonstarter, then, for those who wish to offer up a case that fulfills the demands of S5 but not of S4. The question is how worrisome this really is. It seems like the natural result of involving a thick notion of explanation or priority and thus abandoning the monotonicity which characterizes the standard supervenience analyses. One question is perhaps whether it is still possible, given the collapse of S4 and S5, to hold that the latter holds in virtue of the former. The move from standard supervenience analyses to S4 was at least partly motivated by the belief that correlations, be they contingent or necessary, are not enough. Entailments are not enough to capture the dependency that final value stands in to other things. Once we have gone down route and accepted a thick notion of explanation or priority, perhaps we would be willing to suggest that although S4 is equivalent to S5, there is still a sense in which the former could be prior to the latter.

Before we move on, we should also note what would happen if we were to look at S4, not from the perspective of the broad relation of dependence, but from the perspective of the narrower relation of evaluative resultance. This relation is supposed to involve only the factors that we would judge to be the makers of final value, i.e. the very first things we point to when we want to explain why an object ends up being good or bad for its own sake. Now, it may not seem obvious at first, but this means that although S4 may well be able to adequately capture the dependence of final value, it would probably do less well from the perspective of evaluative resultance. More specifically, the second necessity claim involved in S4 would not hold whenever the nonnormative properties it mentions are understood in a narrow sense, as including only the makers of final value. The reason is that there could be other relations that have an effect, resulting in an object not sharing the final value of another object, even when they are identical in terms of their value-making properties. In other words, then, objects situated on opposite sides, or even the same side, of a modal border, can be identical in terms of their value-makers and still fail to be identical in terms of their final value. These are such crucial points that we should perhaps try to say more about them.

The present issue relates to the question what modal profile of the relations that have so far been our focus. When applied to the broad dependence of final value, S4 is meant to express claims that are conceptually true. The reason is that it is strictly inconceivable that an object should have final value, without there being something else to explain, in a broad sense, why this is so. It is just as inconceivable that two objects should be identical in terms of all their relevant nonnormative properties and yet not share the same final value. The observations made in the previous paragraph entail that things are very different
when we consider evaluative resultance. When applied to this relation, S4 starts to look deeply problematic. Of course, it is also strictly inconceivable that an object should have final value, without there being something to explain, in a narrow sense, why this is so. In this regard, the robust analysis does just fine. However, the second necessity claim involved in S4 will either be false, or it will be true in a sense that is not conceptual. Perhaps it ends up being a first-order issue whether evaluative similarity is preserved in the way suggested by the similarity intuition. We shall return to this issue shortly. What is important here is that there is indeed a strong connection between final value and nonnormative properties, but that it seems that this connection cannot be analyzed in reductive way. Instead, an accurate analysis must involve notions such as explanation and relevancy, themselves hinging on a prior grasp of what is meant to be analyzed, which takes us beyond mere evaluative supervenience.

On the Entailment Dilemma

In this section, we shall stay on the issue of what modal profile should be assigned to the dependence of final value and evaluative resultance, for here there are some potential objections that need to be addressed. Some philosophers find inspiration in the ideas of Hume (1978/1739) and either hypothesize or assume that you cannot derive an ‘ought’ from an ‘is’. More generally, it is often suspected that we cannot derive any normative conclusions from purely nonnormative premises. Blackburn (1971, 1987) is among those philosophers and he argues that if this is right, then there should be a lack of entailment between nonnormative properties and value. This is worrying, because it makes it unclear why it should be necessarily true that two objects that are identical in terms of their relevant nonnormative properties are also identical in terms of their final value. Yet, we have just learned that this is the way things should work: If one object is exactly like another in terms of its nonnormative properties, we can immediately draw the conclusion that it must also be exactly like it in terms of its final value. The challenge set out by Blackburn is not easy to follow, and it has been spelled out in several ways, but within the present context it should be interpreted as a challenge to explain why

112 Often, this is unfortunately introduced to students of philosophy as ‘Hume’s law’ and is summarized with the slogan that “you cannot derive an ‘ought’ from an ‘is’.”
there is this discrepancy in the first place: Whence the guarantee of evaluative similarity across possible worlds, given the lack of entailment Hume warns us of? One possible strategy for dealing with the challenge in question would be to retreat from a robust understanding of dependence and even a strong understanding of evaluative supervenience. As we have seen, a weak understanding of the dependence of final value has it that objects on opposite sides of the border between two possible worlds need not have the same final value even if they are otherwise exactly alike. However, we are then saddled with the inadequacies of weak supervenience. Blackburn (1987, pp. 182–187) puts a dramatic spin on these inadequacies by saying that weak supervenience would leave us without a general explanation for why there is a ban on *mixed worlds*. We could not, then, explain why objects on the same side of a modal border necessarily have the same final value if they are otherwise exactly alike. Most philosophers seem to agree with this sentiment and have therefore formulated other responses to the challenge in question. Another common response is to understand these matters in terms of a kind of necessity that does not run counter to the lack of entailment thesis. For example, Shafer-Landau (2003, pp. 84–85) and Strandberg (2004, pp. 248–249) could be taken to suggest that the ban on mixed worlds is not entirely conceptual, but rather metaphysical and due to the substantive nature of final value.113 It is just the way in which final value behaves that gives us the guarantee of evaluative similarity across possible worlds.

The strategy of relying upon other kinds of necessity in this context seems to me fundamentally mistaken. In fact, the strategy is mistaken for much the same reason that weak supervenience turns out to be far too weak. That two objects should have the same relevant nonnormative properties and yet fail to have the same final value is not just violently counterintuitive from a first-order perspective, but it is utterly inconceivable. Any attempt to show that two objects can be exactly alike in all respects except the evaluative one must assume that the objects are in fact subtly different, in terms of either their intrinsic or their extrinsic properties. This indicates that the ban on mixed worlds is not metaphysical and due to the substantive nature of final value after all. It would be dishonest, then, to pretend otherwise just so that we can attain a straightforward explanation for why the similarity intuition holds true within

---

113 I say that this is roughly their strategy, because they looked at these issues from the perspective of evaluative supervenience and not the thicker sort of relation that have been the focus here. Furthermore, the philosophers in question were concerned with value generally and not just final value. For more on their approach to the challenge set out by Blackburn, see e.g. Dreier (1992, pp. 20–21) and Ridge (2007).
this context. The better strategy for responding to the challenge set out by Blackburn is the more stubborn one of insisting that the second claim involved in the robust understanding of dependence is formally guaranteed.

The insistence is that the concept of final value is such that we simply cannot judge one object to be good for its own sake and another bad for its own sake, unless the objects are also understood to be different in some other respect. While there may be a deeper and more interesting story to tell as to why this is the case, such a story would presumably end up being about why our ordinary concept of final value has the formal features that it has in the first place. Exactly what this story will look like is uncertain, but it is likely to focus on the potential benefits of our first introducing such a concept into our evaluative talk and practices to begin with. It is not hard to predict the disagreements that will be invited by any attempt to formulate such a story in closer detail. Some philosophers would probably argue that to be plausible, the story must make some reference to the benefits of being able to refer to actual final values, understood as genuine properties or relations in the world. Other philosophers would presumably insist that the story can omit such queer features and focus more on the practical and social benefits of talking and thinking about final value in a manner that fits well with the similarity intuition.\footnote{Blackburn (1971, 1985) himself makes a rough attempt at such a story.}

Giving a detailed story as to why value concepts have the formal features that they have goes beyond the purpose of the present work. The same is true for any attempts to resolve the kinds of disagreements that the formulation of such a story might invite. The point here is that it is entirely because of those formal features that we know about the dependence of final value.\footnote{The ban on mixed worlds also follows naturally from the thick notion of explanation or priority which has been appealed to. Should two objects differ in terms of their final value but not in terms of some other set of properties, then it turns out that this set of properties does not explain the final value of the objects after all—at least not in a broad sense. Evaluative resultance is of course more allowing in this regard.} The question is what the cost of adopting this stubborn line of response would be: Does it necessarily result in an outright abandonment of the lack of entailment thesis? Abandoning the lack of entailment thesis might bring with it difficult problems, because it seems to leave the door wide open for reductionism. In the present context, reductionism should be understood as the view that final value must be identified with the nonnormative properties on which it depends. The thought is that if the presence of final value entails the presence of nonnormative properties and the presence of some nonnormative properties entail the presence
of final value, this indicates that the former can be reduced to the latter. ¹¹⁶ We shall return to this issue in just a moment.

For now, we can just insist that it is not incompatible with any reasonable interpretation of the lack of entailment thesis to claim that two objects could not differ only in terms of their final value. It is just that without the entailment from nonnormative properties to final value, we must give up one way of explaining why the claim is true. This means that no reasonable interpretation of the lack of entailment thesis stands in direct opposition to S₄, although perhaps it could be understood as undermining the intuition of similarity. One interpretation of the lack of entailment thesis suggests that we cannot conceptually infer what final value an object has based purely on observations about its nonnormative properties. This seems to me to capture the spirit of the tradition rooted in the ideas of Hume. This modest version of the lack of entailment thesis is also perfectly compatible with the claim that has just been made, that we can nonetheless conceptually infer that two objects that are identical in terms of their relevant nonnormative properties must share the same final value. If this view also ends up being rejected by some of the more vociferous followers of Hume, then I think that we should have no qualms whatever about refusing to join their ranks. ¹¹⁷

What of the question of reductionism? Would the identity of final value with nonnormative properties necessarily follow from S₄? The correct answer appears to be a negative one, for there is seemingly nothing in the analysis itself that commits us to postulating an identity between final value and the nonnormative properties on which it depends. In fact, that the dependence in question appears to have an asymmetric and irreflexive character would even seem to exclude the possibility of such an identity. However, here things start to become a bit

¹¹⁶ For an interesting criticism of the objection, see Oddie (2005, pp. 146–175).
¹¹⁷ Some of the things that have been said here might raise suspicion that, like Blackburn, we will be assuming that descriptivism cannot be true. This suspicion is unfounded. We should want our robust analysis of dependence to be compatible with a reasonable interpretation of the lack of entailment thesis, which it fortunately appears to be, without committing ourselves to it. The reason why we want it to be so compatible is that to date, no proposed reduction of final value appears to hold true. No specific nonnormative properties have so far appeared to conceptually necessitate the presence of final value. We should be open to the possibility, however, that there may be a plausible reduction waiting to be discovered. Whether this ends up happening or not, i.e. whether the lack of entailment thesis ends up being true or false, will have little relevance with respect to the plausibility of the robust analysis of dependence, or our understanding of evaluative resultance. More will be said about these issues in the upcoming chapters, as we will frequently tread on the line of appearing to beg the question against reductionism, but without actually doing so. Let us now move on from the lack of entailment worry and briefly say something about the varieties of dependence bases that there are.
complicated. It is true that the dependence in question appears to have an asymmetric and irreflexive character. It is also true that final value is multiply realizable in the sense that it can conceivably depend upon many different nonnormative properties. Nevertheless, we could in principle list all the possible dependence bases for final value in an infinite disjunctive description. Such a description, when treated as a predicate, would be necessarily coextensive with the concept of final value. According to Jackson (1998), it follows that the two must pick out the very same property:

Now each world at which \(E\) is true will have some descriptive nature: ethical nature without descriptive nature is impossible (an evil act, for example, must involve death or pain or...). And, for each such world, there will be a sentence containing only descriptive terms that gives that nature in full. Now let \(w_1, w_2, \ldots\) be the worlds where \(E\) is true, and let \(D_1, D_2, \ldots\) be purely descriptive sentences true at \(w_1, w_2, \ldots\), respectively, which give the full descriptive nature of \(w_1, w_2, \ldots\). Then the disjunction of \(D_1, D_2, \ldots\), will also be a purely descriptive sentence, call it \(D\). But then \(E\) entails and is entailed by \(D\) (1998, pp. 122–123).

118 Although it is of relevance to any area of potential reductionism, the notion of multiple realizability was originally introduced in the philosophy of mind by Putnam (1967). For a helpful overview of sources and discussions surrounding this concept, see Bickle (2016).

119 The general possibility of this argument was first pointed out by Kim (1984). Its applicability to the case of value was then mentioned by Depaup (1987), before it was articulated in detail by Jackson (1998). There have been worries that the argument generalizes too much to all dependent phenomena, but some philosophers, like Streumer (2008), have argued that these worries are overstated. See also Brown (2011) for a more recent defense.

120 Although Jackson here has in mind ethical sentences and formulates the argument in terms of global supervenience, the same argument is relevant to goodness and badness more generally, as well as the dependence of final value. For some technical complications having to do with the understanding of nonnormative properties that underlies the argument, see Dunaway (2015, pp. 631–632). For a selection of papers dealing with the precise nature of the property that would be picked out by the infinite disjunctive description on which Jackson relies, and specifically his broader meta-ethical theory of moral functionalism, see e.g. Jackson & Pettit (1995, 1996), Jackson, (1998), Van Roojen (1996), Yablo (2000), Zangwill (2000), Williamson (2001), Horgan & Timmons (2009, and Streumer (2011). Arguments that run parallel to the one advanced by Jackson in this section have also been debated in other areas of philosophy. They occur in discussions concerning determinable properties (Rodriguez-Pereyra, 2002), truth (Pedersen, 2005), color (Ross, 2010), and, perhaps most notably, mind (where it is often referred to as Kim’s supervenience argument) (e.g. Block & Fodor, 1972; Fodor, 1974; Owens, 1989; Seager, 1991; Kim, 1984, 1990, 1992, 1998; Marras, 1993; Zangwill, 1995;
One question is whether it is true that necessary coextension entails identity in this way. Another question is whether the infinite disjunctive predicate appealed to here picks out a genuine property. If the correct answer to either of these questions is a negative one, this would severely undermine the argument just considered. The next sub-section considers these questions in turn. First, we will look at strategies that aim to show that the disjunctive predicate appealed to here does not pick out a genuine property. This strategy is usually coupled with attempts to formulate some sort of requirement or condition that must be fulfilled for any predicate to be property-preserving to begin with. We will then move on to strategies that aim to show that the identity of properties cannot be inferred from the necessary coextension of the predicates that pick them out. The fact that two descriptions are necessarily coextensive, then, does not mean that they ascribe the same properties to objects. An argument will be made in favor of the second strategy over the first, before we move on to some other more general worries about the nature of dependence.

On Property-Preserving Predicates

The first question to be considered is whether the infinitely disjunctive predicate appealed to by Jackson picks out a genuine property. Streumer (2008, p. 546) rightly points out that a negative answer to the question cannot plausibly be motivated by the belief that disjunctive predicates generally are incapable of picking out properties, for this is clearly false. If there is a property of having an additive primary color, he says, then this property can be picked out by the disjunctive predicate ‘is blue or red or yellow’. Streumer (2008, p. 547) also suggests that a negative answer to the question cannot plausibly be motivated by the belief that infinitely disjunctive predicates generally are incapable of picking out genuine properties, for according to Streumer this is false as well. If there is a property of having color, he says, then this property can be picked out by an infinite disjunctive predicate which lists all the possible colors that there are. A less general version of the skeptical strategy would have to be devised.

Clapp, 2001; Jaworski, 2002; Gozzano, 2003; Walter, 2003, 2006). For an argument to the effect that global supervenience is not enough to generate the reduction argument, but something like strong supervenience needs to be used, see Williamson (2001).

If one is skeptical that there is an infinite amount of colors, there are less problematic examples that might be used. Wlodek Rabinowicz has suggested to me (in private communication) that one look at the class of natural numbers.

\[121\]
Obviously, we cannot settle for simply insisting that there is something particularly strange about the predicate that Jackson has in mind, for this would beg the question against his views.

One criterion on property preservation states that predicates pick out genuine properties if, and only if, the properties have causal powers.\(^\text{122}\) The intuition is that for a property to be instantiated is for it to make a difference and that this should be understood in causal terms. The argument, then, is that the infinitely disjunctive predicate relied upon by Jackson is particularly strange, because the property it picks out cannot possibly make a difference in this sense.\(^\text{123}\) Jackson (1998, pp. 105–108) is aware of the objection and does not find it very convincing. That a painting features mainly additive primary colors might cause a person to buy it for her house. If this is right, then the property picked out by the description ‘is blue or red or yellow’ does have causal powers. Examples similar to this motivate Jackson to say that complex properties, like those picked out by infinitely disjunctive predicates, may indeed be causes, so long as they are not too disjointed. Of course, to show that the property picked out by his infinitely disjunctive predicate is not too disjointed, Jackson has to show that it displays enough uniformity or cohesiveness. Whether he can really show this seems to me rather doubtful.\(^\text{124}\)

More could be said about the causal criterion, but we shall leave it aside for now. One reason is that it is difficult to apply and gain argumentative leverage from the causal criterion in the present context. To me the causal criterion seems to pose no difficulty for the antirealist about final value, who might welcome the conclusion that the infinitely disjunctive predicate on which Jackson relies does not pick out a genuine property. In other words, this is not a conclusion that would motivate her to give up on the reduction that Jackson proposes. The causal criterion also poses no difficulty for the realist about final value, who might just reject it and welcome the conclusion that final value is itself causally ineffectacious. Another reason for moving away from the causal criterion is that it might appear far too strong as it is. There are arguably plenty of abstract properties that lack causal powers. At the very least, then, it would be

\(^{122}\) For an influential defense of the causal criterion, see Shoemaker (1984, chapters 10–11).

\(^{123}\) Non-reductive realism is frequently seen as a mysterious view. One reason is that it understands value as being dependent upon, but nonetheless completely distinct from, everything else. Much of the allure of descriptivism is presumably due to its avoiding this result and instead placing value in the causal world. The worry just raised is that this benefit might be jeopardized by the views of Jackson.

\(^{124}\) The issue seems to partly hinge on the plausibility of his controversial meta-ethical theory of *functionalism*. For sources on the subject, see footnote 120.
nice if we could formulate a criterion regarding property preservation that was somewhat less entrenched in this particular kind of metaphysics.

By drawing on the work done by Gärdnfor (2000, 2014) on concepts and cognitive representations, Oddie (2005, pp. 152–158) tries to give us an alternative criterion for determining when predicates end up being property-preserving. The criterion he has in mind says that a predicate picks out an ordinary property if, and only if, it can be modelled as a convex region in a conceptual space. A conceptual space should here be understood as a geometric model representing properties or sets of properties. A simple one-dimensional temperature space that he uses as an example models the properties of being cold, warm, and hot (see Fig. 2 below). For illustration, we should think of these properties as being mutually exclusive and collectively exhaustive of the relevant temperature space. Within this conceptual space, the property of being cold or warm, on the one hand, and the property of being warm or hot, on the other hand, both turn out to be ordinary properties. The property of being cold or hot, by contrast, is not an ordinary property. The reason is that the property of being cold and the property of being hot, are separated by a non-constituent, namely the property of being warm.

The example is meant to demonstrate that ordinary properties, i.e. the kind of properties that tend to figure in our everyday talk and practices, carve out convex regions in conceptual spaces, meaning that their constitutive points are not separated by any nonconstitutive points. In slightly more technical terms, a convex region is such that for any two of its points, x and y, every point between x and y (e.g. every point on the straight line between x and y) also belongs to this region. This clarifies the sense in which the convexity-criterion is naturally

---

125 For more on philosophical applications of conceptual spaces, see Zenker & Gärdenfors (2015).
126 Oddie, like Gärdnfor, uses the term ‘natural’, but here we will instead use the term ‘ordinary’, to avoid any confusion. By using the term ‘ordinary’, it is made more obvious that we do not take this to mean the same thing as ‘naturalistic’ or ‘nonnormative’.
127 This property can obviously be picked out by means of negation as well. The property of being cold or hot is captured by the description ‘is not warm’. Oddie (2005, p. 154) means that regardless of how it is picked out, the property of cold or hot fails to take up a convex region in the temperature space example.
underpinned by intuitions of *betweenness* (e.g. Gärdenfors, 2000, pp. 15–17, 69–80; 2014, pp. 24–29). For the sake of illustration, we can imagine a color space modeling all the nuances that are visible to the human eye. The color space will then consist in tens of thousands of points, ranging from the paradigmatically blue, red, yellow, and every subtle variation between them. It seems natural, if not obvious, that if two randomly selected points in the color space both find themselves within the blue region, then any points between them would also find themselves within the blue region. This clarifies how the blue would be modeled in accordance with the convexity-criterion.

After illustrating the convexity-criterion further, Oddie (2005, pp. 158–162) turns his attention to the case of value. He suggests that while value can be modeled as convex regions in normative conceptual spaces, value will correspond to nonconvex regions in nonnormative conceptual spaces. He also argues that when value and its nonnormative basis are modeled in the same conceptual space, “convex conditions at the level of value will be highly disjointed at the level of the natural” (2005, p. 162). Oddie thinks this indicates that the former cannot be reduced to the latter. If this is on the right lines, Jackson and his sympathizers need to do more than show that the concept of final value is necessarily coextensive with an infinite disjunctive predicate, for such a predicate may not pick out an ordinary property. Oddie could stop here and settle for offering a challenge to reductionism and its proponents. Admirably, he (2005, pp. 152–179) tries to develop his argument further and demonstrate that some of the properties picked out by disjunctive predicates cannot be plausibly modeled as a convex region in the way just described. What follows is a brief overview of his main points, proceeding from some of the simpler two-dimensional conceptual spaces that he makes use of:

---

128 Oddie (2005, pp. 160–162) mentions the example of cruelty and its respective determinates mild cruelty and severe cruelty. Frustrating cats for fun has the property of being mildly cruel, he says, while subjecting cats to torture has the property of being severely cruel. What the convexity criterion implies is that if frustrating cats for fun and torturing cats for fun are determinates of cruelty and have the property of being morally bad, the same must be true for the act of tormenting cats. The reason is that it lies between the points occupied by the property of being mildly cruel and the property of being severely cruel. However, Oddie thinks that when we look at cruelty more generally and consider the nonnormative bases that might account for its negative value, things do not end up looking so neat. The nonnormative bases that account for the badness of cruelty can be shown to be quite shapeless. We will return to the concept of shapelessness at several points in the discussion to follow.
Fig. 2 is meant to model the states that two hypothetical people, x and y, can be in at a given time. On the horizontal axis (in non-italics), we see a modeling of the states of x, ranging from the state of pain to the state of neutrality to the state of pleasure. On the vertical axis (in italics), we see a modeling of the states of y, ranging from the state of desire to the state of indifference to the state of aversion. It is important that we here understand the states of y as being about the states experienced by x, but not the other way around. For example, in the bottom right corner, we see a point, 9a, which represents the state of y feeling aversion toward the pleasure experienced by x. By contrast, in the top left corner, we see a point, 1a, which represents the state of y feeling desire toward the pain experienced by x. Suppose we rank all the points in terms of the value of y’s responses to x’s states, from best to worst. Points 1a and 9a may well be regarded as the worst. The state of y feeling aversion toward the pleasure of x and the state of y feeling desire toward the pain of x both appear very bad indeed. When taken together, we might suppose that they constitute the dependence base for the complete impropriety, and hence badness, of y’s responses regarding the states of x. However, Oddie (2005, pp. 164–166) notes that none of the points figuring between the two points in question, 2a–8a, will also constitute the dependence base for the complete impropriety of y’s
responses regarding the states of x. For example, the state that y is indifferent about the neutral state of x (as in the point 5a) is not completely inappropriate, but rather the opposite. In this case the attitude adopted by y regarding the states of x is completely fitting.

Consider next the point in the top right corner, 3a, which is meant to represent the state of y desiring the state of pleasure that is being experienced by x. The bottom left corner features a point, 7a, which represents the state of y feeling aversion toward the pain of x. In the middle of the conceptual space, there is a point, 5a, which represents the state of y being indifferent regarding the neutral state that is being experienced by x. On the ranking we have made, all three may well be the best. According to Oddie (2005, pp. 164–166), they all correspond to the complete propriety, and hence goodness, of y’s responses regarding the states of x. Once again, when 7a, 5a, and 3a are taken together, the three points make up a condition that is not convex. Between 7a, 5a, and 3a, there are several points, namely 2a, 6a, 4a, and 8a, which do not correspond to the complete propriety of y’s responses regarding the states of x. For example, the point 8a represents the state of y being averse to the neutrality that is experienced by x. This point may not represent a state that is completely inappropriate and hence bad, but neither is it completely appropriate and hence good. In terms of value, it should instead find itself somewhere in the middle of our overall ranking. According to Oddie, this indicates that the predicates ‘is completely appropriate’ and ‘is completely inappropriate’, although themselves subject to convex treatments, do not correspond to nonnormative conditions that are convex. Consider next the following:

For the sake of this specific conceptual spaces, Oddie uses a city-block metric to cash out the definition of betweenness that was just explicated. This means that you can travel between points either horizontally or vertically (i.e. along the dotted and non-dotted lines), but not along the diagonal. It is of course true that the extent to which a region displays convexity will depend upon the metrics used to cash out the definition of betweenness (cf. Gärdenfors, 2000, p. 18) and perhaps it is also true that different sorts of conceptual spaces will call for different sorts of metrics. There is a lot of relativity in play here. We shall return to this point later but will for now treat it as a relatively unproblematic aspect of Oddie’s argument.

The argument just presented relies upon a holistic view of the structure and behavior of value (Oddie, 2005, p. 168). In this context, the view states that different objects can contribute value to complex wholes in a surprising way that does not reflect the sum of the values that the objects have individually. The state of someone experiencing pleasure might be good for its own sake, while the state of indifference might be neutral regarding its final value. Yet, the combination of someone being indifferent toward the state of someone experiencing pleasure might be bad for its own sake. We return to the issue of evaluative holism later.

---

129 For the sake of this specific conceptual spaces, Oddie uses a city-block metric to cash out the definition of betweenness that was just explicated. This means that you can travel between points either horizontally or vertically (i.e. along the dotted and non-dotted lines), but not along the diagonal. It is of course true that the extent to which a region displays convexity will depend upon the metrics used to cash out the definition of betweenness (cf. Gärdenfors, 2000, p. 18) and perhaps it is also true that different sorts of conceptual spaces will call for different sorts of metrics. There is a lot of relativity in play here. We shall return to this point later but will for now treat it as a relatively unproblematic aspect of Oddie’s argument.

130 The argument just presented relies upon a holistic view of the structure and behavior of value (Oddie, 2005, p. 168). In this context, the view states that different objects can contribute value to complex wholes in a surprising way that does not reflect the sum of the values that the objects have individually. The state of someone experiencing pleasure might be good for its own sake, while the state of indifference might be neutral regarding its final value. Yet, the combination of someone being indifferent toward the state of someone experiencing pleasure might be bad for its own sake. We return to the issue of evaluative holism later.
The conceptual space above incorporates both evaluative and nonnormative aspects, concerning the states of x and the relative propriety of y’s responses to said states. On the horizontal axis (in non-italics), we once again see a modeling of the states of x, ranging from the state of pain to the state of neutrality to the state of pleasure. On the vertical axis (in italics), we this time see a modeling of the relative propriety of the responses of y, ranging from completely inappropriate to somewhat appropriate/inappropriate to completely inappropriate. It is still important that we here understand the states of y as being about the states of x, but not the other way around. Oddie (2005, pp. 170–171) points out that value predicates like ‘is good’ here carve out perfectly convex regions and, so long as we have an attractive first-order theory at hand, the regions in Fig. 3 should be relatively easy to map onto the regions in Fig. 2. For example, the predicate ‘is good’ carves out the region made up of points 1b–

---

131 Włodek Rabinowicz has pointed out to me (in private communication) that there might be something counterintuitive about this combined space from an evaluative standpoint. One problem is that point 1a is equidistant to 3a and 7a. Two points of equal value should be closer together than two points with radically different values. Another problem is that it is unclear how we should map 8b to the previous space.
3b in Fig. 3, but it corresponds to the region made up of points 7a, 5a, and 3a in Fig. 2. The predicate ‘is bad’ carves out the region made up of the points 7b–9b in Fig 3, but it corresponds to a region made up of points 1a and 9a in Fig. 2. The region made up of points 4b–6b in Fig. 3, picked out by the predicate ‘is somewhat bad/good’, maps onto the points 2a, 4a, 6a, and 8a in Fig 2.\(^{132}\)

Aside from expressing some general misgivings regarding the details of the convexity criterion, and thus the general framework of conceptual spaces, Streumer (2008, p. 546, n. 20) argues that the view on which Oddie relies is vulnerable to counterexamples. To illustrate this, he mentions a case involving additive primary colors. Whereas the predicate ‘is an additive primary color’ can be modeled as a convex region in a conceptual space, the same is not true for the description ‘is blue or red or yellow’. According to Streumer, the reason for this is that the region carved out by the description ‘is blue or red or green’ has subregions between which we will find the things carved out by the predicates ‘is yellow’, ‘is magenta’, and ‘is cyan’. The point here seems to be that if we follow Oddie’s approach, then we would need to deny that being an additive primary color is also reducible to being red or blue or green, but this would be absurd. Perhaps there is something to this. Think back to the simple temperature space with which we began. We can imagine an expansion of this conceptual space so that it included all the temperatures that human beings might come across in their daily lives. We can then construct a disjunctive predicate that lists all the points in the space which are occupied by what would be considered

---

\(^{132}\) The objection could be made that whether a predicate ends up carving up a convex region or not depends upon how we design the relevant conceptual space and possibly what metric is used when we consider the distances and relations between its points (cf. Gärdenfors, 2000, pp. 17–20). It is not surprising, then, that what is at bottom the same property can be shown to be convex relative to one conceptual space and non-convex relative to another. If Oddie wants to be convincing, he needs to model the relevant evaluative and nonnormative properties within the same conceptual space. On the other hand, if he were to do so and have them occupy different regions altogether, this would be tantamount to begging the question against reductionism, for it would be to treat the relevant properties as if they were indeed not identical. Interestingly, Oddie (2005, pp. 172–180) later incorporates both Fig. 2 and Fig. 3 as subspaces within the same overarching and three-dimensional conceptual space. The purpose of this more complex conceptual space is to incorporate all the evaluative and nonnormative aspects which have been assumed to be relevant to the case at hand. This obviously means that Oddie can no longer show that a predicate like ‘is good’ picks out a convex evaluative region and a non-convex nonnormative region. Nevertheless, he still manages to find another route to the conclusion he wants to reach, by considering the question which subspace affords regions with their convexity. Oddie shows that predicates like ‘is good’ can be projected upon convex regions in the evaluative subspace and non-convex regions in the nonnormative subspace. Such a predicate may therefore carve out a region that is “convex in virtue of the geometry of value, not in virtue of the geometry of nature” (2005, p. 176).
uncomfortable temperatures (by saying that a temperature is uncomfortable I mean to say that it tends to elicit discomfort and nothing else). The region that is constituted by these points would not be convex, since it excludes the points found in the middle of the space. However, it seems as if it might be easy to construct an alternative conceptual space in which the property that is picked out by the predicate ‘is an uncomfortable temperature’ would indeed be convex.

Although these kinds of counterexamples are perhaps not conclusive, they may cause us to have reasonable doubts about the convexity criterion, at least as it is here used. Oddie (2005, pp. 154–155) rightly points out that the framework of conceptual spaces has been shown to be fruitful, and that the convexity criterion has gained a significant amount of empirical support. However, it would be a mistake to draw the conclusion that the convexity criterion works as a general metaphysical criterion for when something counts as a genuine property. It works well for certain kinds of properties, especially perceptual ones that are expressed through simple words in natural language, but it is likely to do less well from the point of view of more abstract concepts that require language to be learned. In the end, I think it is unclear what the argument by Oddie really shows. Perhaps it is best seen as a rough attempt to illustrate a more basic intuition which we will also be discussing in the next section. What Oddie demonstrates is primarily that the nonnormative basis for value seems marked by a kind of shapelessness when it is modeled in a conceptual space. However, just like it is a mistake to think that a property is shapeless because the predicates used to pick it out are shapeless, perhaps it is also a mistake to think that a property is shapeless because it may be shapelessly represented in a purposefully relativistic model of cognitive representation, whose purpose is not to settle any issues about the metaphysics of properties, but to help make empirical predictions about learning and behavior.

From Necessary Coextension to Identity

alleged counterexamples that they rely upon are criticized by the likes of Jackson (1998, pp. 125–128), Streumer (2008, pp. 541–545), Oddie (2005, pp. 151–152), and Brown (2011, pp. 215–216). The most common rejoinder among critics is that the alleged counterexamples do not involve cases of necessary coextension to begin with, or that they must involve cases of identity. An intuitive standoff has resulted, with neither side being very likely to change their minds unless further arguments are presented by the opposing side. Brown (2011, pp. 212–213) and Enoch (2011, pp. 139–140) offer a diagnosis of the situation, on which philosophers who deny the legitimacy of the alleged counterexamples do so mainly out of concerns for ontological parsimony. Perhaps there is some truth to this diagnosis. We shall return to it below but let us first take a closer look at a few of the alleged counterexamples that have emerged from the literature so far, so that we can get a firmer grasp of the disagreement. Streumer (2008, p. 542) at one point considers the classic pair of descriptions ‘has three sides’ and ‘has three angles’. 134 He concedes that these descriptions pick out different properties but points out that they are not necessarily coextensive. The reason is that there is a geometrical figure that satisfies the former description without satisfying the latter. The example given by Streumer involves a box-shape or an incomplete square with one of its four sides missing. Such a geometric figure, he points out, would have three sides but only two angles. Streumer (2008, p. 543) therefore moves on to the pair of descriptions ‘is a closed figure with three sides’ and ‘is a closed figure with three angles’. He concedes that these descriptions are in fact necessarily coextensive, but like Jackson (1998, p. 126) before him, Streumer also insists that they must pick out the very same property. 135 According to Streumer, the pair of descriptions ‘is a closed figure with three sides’ and ‘is a closed figure with three angles’ pick out the property of objects having the following shape:

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure}
\caption{A geometrical figure with three sides but only two angles.}
\end{figure}

\begin{table}
\centering
\begin{tabular}{|c|c|}
\hline
Property & Example \\
\hline
Three sides & Three angles \\
\hline
\end{tabular}
\caption{A table showing the necessity of coextension.}
\end{table}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure}
\caption{A geometrical figure with both three sides and three angles.}
\end{figure}

134 See also Sober (1982).
135 It may be insisted that it is possible to know or determine that a closed object has three angles without knowing that it has three sides, and vice versa. I am not entirely sure what to say to this argument. One potential problem with examples of this kind is that it can always be insisted that one and the same property has different epistemic modes in the sense that it can figure in our thoughts in separate ways (Jackson, 1998, p. 126).
Another counterexample that is frequently brought up in the literature and which is criticized by Streumer (2008, pp. 544–545) concerns non-geometric mathematical descriptions. The pair of descriptions ‘is the number two’ and ‘is an even prime’ are necessarily coextensive, but that they should also pick out the same property seems controversial. Streumer responds by first pointing out the fact that mathematical descriptions such as these seldom involve claims of predication to begin with but are instead used to make claims of identity. The claim ‘two is an even prime’ is one among countless examples of such an identity claim. This means that the descriptions are such that they do not have the required form to pick out properties. Streumer furthermore suggests that even in contexts where they do involve claims of predication, it seems reasonable to assume that the descriptions imply an identity (Streumer, 2008, p. 545).

If we are concerned with parsimony and do not want to accept more properties than are indicated by the extensions of our predicates, then perhaps we want to join Streumer and his sympathizers in their heel digging. However, there may appear to be little in the way of argument that could persuade us to do so, should our concerns for parsimony give way to substantive intuition. Suikkanen (2010) tries to break the stalemate that has resulted from these

---

136 Kramer (2009, pp. 210–211) thinks that Streumer is too quick to reject the first example just mentioned, involving the pair of descriptions ‘has three sides’ and ‘has three angles’. It is true that if we consider them in an open context, the descriptions in question are not necessarily coextensive. However, things start to look different if we specify that we are considering the pair of descriptions from the point of view of closed geometric figures. For example, when the descriptions ‘has three sides’ and ‘has three angles’ are applied to triangles, they are also necessarily coextensive, but does this mean that they pick out the same property of triangles? This response is unfortunately not very convincing. Streumer is just as likely to say that in such narrow contexts, the pair of descriptions ‘has three sides’ and ‘has three angles’ do pick the same property. They pick out the same property that is picked out by the descriptions ‘is a closed figure with three sides’ and ‘is a closed figure with three angles’. There is still a standoff here that seems unlikely to be overcome by the formulation of further counterexamples.
discussions by appealing to Leibniz Law, which states that two things are identical if, and only if, they have all the same properties. The law provides a test for identity, since it entails that if two things do not have all the same properties, then they cannot be identical. Among other things, Suikkanen wants to show that the higher-order properties of final value are different from the higher-order properties of the feature that is picked out by the kind of infinitely disjunctive predicate on which Jackson relies. Suikkanen (2010, pp. 99–104) mentions several candidates for what these higher-order properties might be, among them being the sort of property we looked at in the earlier subsection. Value differs from its nonnormative basis in so far as he latter is marked by a kind of shapelessness. There is an intuitive sense in which nonnormative properties lack the kind of unity which is displayed by value. Suikkanen is skeptical of this suggestion, for much the same reason that we dismissed it earlier. However, his skepticism is formulated against the background of a debate in metaphysics about the status of disjunctive properties.

D. M. Armstrong (1978, pp. 19–23) famously argues that disjunctive properties are not real, for they do not mark genuine similarities between distinct objects. To say that one object that has the property of being round is similar to another object that has the property of being red in respect of their common property of being round or red seems absurd. However, the problem with this sort of argument is that the notion of a property being disjunctive is a bit misleading. As Oddie (2005, p. 151) and Streumer (2008, p. 545) independently point out, what is disjunctive is just the predicates used to pick out the property in question. Whether the property itself is disjunctive is an open question that is, quite frankly, difficult to even understand. This makes it seem as if any criterion that tries to rule out the property-preserving nature of the description on which Jackson relies cannot plausibly find support in skepticism regarding disjunctive properties as such. Suikkanen agrees that criteria of this kind can also not rely too heavily on observations about the shapelessness that marks our predicates. This observation resonates well with my earlier observations about the argument from conceptual spaces.

Like Jackson, Suikkanen is not explicitly interested in final value but in rightness. Once again, this is a difference that makes no difference. The same arguments apply in this context.

Of course, proponents of disjunctive properties might not see the conclusion as absurd but would claim that any two objects that have ordinary properties will also have disjunctive properties which mark a similarity between the objects. For some recent sources and overviews of these kinds of discussions, see e.g. Audi (2013) and Skiles (2016). Because of the reasons highlighted by Streumer above, we will grant that some disjunctive descriptions can pick out genuine properties and that there are, in this technical sense, disjunctive properties.
Suikkanen (2010, p. 101) thinks that the way we attain knowledge about value is different from the way we attain knowledge of the nonnormative property picked out by the infinite disjunctive predicate on which Jackson relies. Value therefore has a different epistemic profile from its nonnormative basis. This is intuitive at first, but it seems to me to be vulnerable to a similar objection to the one which has just been mentioned. Parallels can be drawn to disagreements within the philosophy of mind. In this context, Jackson (1982, 1986) is famous for arguing that we attain knowledge of what it is like to experience red in a way that is fundamentally different from the way we attain knowledge of the physical processes that may be correlated with experiences of red. At one point, this tempts him toward the conclusion that experiences of red cannot be identified with the physical processes in question. His arguments on this point have been met with countless objections, some of which state that the experience of red can be represented to us in diverse ways (Nida-Rümelin, 2015, section 4.6). Thus, it isn’t obvious that this experience cannot be the same thing as the physical processes.139

It is arguable, then, that just like one and the same thing could have different and irreducible conceptual-space representations, one and the same thing could have different and irreducible epistemic representations. There is an argumentative gap here which needs to be traversed. This kind of response is perhaps less convincing when we consider some of the other higher-order properties that Suikkanen brings up in his discussions. Normative properties, as we have seen, are necessarily dependent upon other things. If an object has final value, then there must be something else that explains, in a broad and narrow sense, why this is so. It is not obvious that this is also true of the property that is picked out by the infinitely disjunctive predicate relied upon by Jackson. It might be argued that the property picked out by this predicate depends upon the properties picked out by its disjuncts. Suikkanen (2010, pp. 101–102) expresses doubt about this suggestion, for he sees this as a straightforwardly

---

139 The most famous dramatization that Jackson (1982, 1986) presents of this argument involves a scientist named Mary, who grows up in a closed room that is completely devoid of color. However, Mary learns everything there is to know about the physical processes involved in color vision. She has access to all present and future research into this subject. Jackson then has someone opens the door and throws in a bright red object. The question is whether Mary learns anything new when she sees the red object and has the experience for herself. Jackson suggests that if she does, then this should mean, roughly, that there is more to the experience of red than the physical processes which underlies it. In the philosophy of mind, this has become known as the knowledge argument. For a collection of important papers on the argument, with additional sources and criticisms, see Ludlow, Nagasawa & Stoljar (2004).
logical relationship, rather than a case of the thick notion of priority or explanation. I do not think that this response is very convincing, for, as we shall soon see, there is an argument to be made that in addition to the straightforwardly logical relation that holds between disjunctions and disjuncts, perhaps it also makes sense to speak of a dependence relation here.

The last difference that Suikkanen (2010, pp. 102–104) mentions in his discussions is more nebulous, but it is one which we have had occasion to consider ourselves in previous chapters, and I think it is quite convincing. The last difference concerns the fact that final value is an inherently normative property in the sense that it is tightly connected to the domain of reasons, but the nonnormative property picked out by the infinitely disjunctive predicate relied upon by Jackson appears not to be. In other words, the property that Jackson has in mind does not appear to be such that it inherently calls upon us to adopt certain responses. In comparison to final value, it is utterly normatively inert. In fact, as Suikkanen (2010, p. 103) points out, the same can be said in comparison to reasons themselves. Reasons have the property of being such that they tug at us and recommend that we adopt certain responses. Again, the nonnormative property which is picked out by the infinitely disjunctive property relied upon by Jackson does not. This is the way things seem, at least, and when it is taken together with the alleged counterexamples mentioned earlier in this section, it seems to me to cast significant doubt on the entailment from necessary coextension to identity. The principle stating that necessary coextensive predicates must pick out identical properties requires additional defense, or else it can be rejected as a general metaphysical principle.

Remarks on the Concept of Grounding

The dependence of final value shares structural features with evaluative supervenience, but it differs from it in the sense that it involves a thick notion of explanation or priority. The dependence of final value upon nonnormative properties, much like the evaluative resultance that it subsumes, imposes a partial order on the domain. It is a transitive, irreflexive, and asymmetric relation. It is also marked by a distinctly hyperintensional flavor, given its resistance to wholly intensional accounts, like those formulated in terms of necessary correlations. Comparisons can therefore be drawn to more general
debates in metaphysics, wherein the concept of *grounding* has become one of the major points of controversy in recent years.\textsuperscript{140} Just like the dependence of final value upon nonnormative properties, grounding is closely tied to metaphysical explanation and priority. What is more, it is also described as a partial order and as having a hyperintensional flavor.\textsuperscript{141} It is therefore tempting to say that the relation we have had in mind just is grounding as it occurs in the normative domain.\textsuperscript{142} The dependence of final value upon nonnormative properties would presumably correspond to the full grounding of final value, while the evaluative resulstance that it subsumes would correspond to the partial grounding of final value. Philosophers often introduce grounding by giving a list of substantive examples: Disjunctions are said to be true because at least one of their respective disjuncts are true; the density of an object is assumed to be grounded in its mass and volume; couples are taken to be married because their respective members are married to each other; cups are brittle because of the way their constituent atoms are arranged; propositions are said to be true because of the way the world

\textsuperscript{140} Among the philosophers most responsible for the recent interest and controversy concerning grounding are Fine (2001, 2010, 2012), Rosen (2010), and Schaffer (2009, 2012). Since then there has been an avalanche of papers on the subject that is difficult to keep abreast of.

\textsuperscript{141} Grounding is sometimes treated as an operator rather than a predicate, along the lines proposed by Fine (2012, pp. 43–46). This is supposed to mean that we need not take a stance on the reality of the things that are meant to stand in the grounding relation. Indeed, we may even have doubts regarding the ontological status of the grounding relation itself. For more on operational or predicative understandings, see Correia & Schnieder (2012, pp. 10–12).

\textsuperscript{142} One complication is that grounding is often assumed to be propositional in nature, meaning that it is meant to hold between propositions, facts, and entities of a similar structure (e.g. Rosen, 2010; Audi, 2012). By contrast, we are assuming that normative explanations may appeal to relations between all kinds of things, including concrete objects, states of affairs, facts, relations, properties, structures, events, and so on. We have the options of either challenging the consensus on grounding or assuming that the dependence of final value is a different kind of relation. My preference is for the former, since the relations appear identical in all other respects. Should readers not share my preference in this regard, that is fine. The arguments and objections we will be considering in this section would apply to grounding and the dependence of final value alike, irrespective of whether they are seen as identical or merely similar relations. Even when it is compared to philosophical discussions at large, the grounding literature appears like the Wild West in that little has so far been settled and everything is up for grabs. For challenges about the propositional nature of grounding, see e.g. Schaffer (2010). For challenges to the irreflexivity of grounding, see e.g. Fine (2010) and Krämer (2013). For challenges to the asymmetry of grounding, see e.g. Rodriguez-Pereyra (2005). For challenges to the transitivity of grounding, see e.g. Schaffer (2012). For additional sources and examples to illustrate these challenges, see Bliss & Trogdon (2016). Finally, for an impressive discussion of several of these topics, as well as issues about whether metaphysical grounding differs in some important respect from normative grounding, see Berker (forthcoming).
happens to be; nonempty sets exist because their elements exist; mental states are thought to be instantiated because brain states are instantiated; and so on\textsuperscript{143}

After listing these kinds of substantive examples, philosophers often contrast metaphysical explanations involving grounding with causal explanations. Grounding supports metaphysical explanations much like causal relations support causal explanation (Smid, 2017, p. 72). Here, my suspicion is that substantive examples from the normative domain might be particularly helpful. A painting can be good for its own sake in virtue of its colors and shapes. The property of being good for its own sake and the property of having certain colors and shapes are both had by the painting at the same time. It is not, then, as if the presence of its colors and shapes causes the painting to be good for its own sake (cf. Audi, 2012, p. 687). Nor need we assume that the painting being good for its own sake is identical to, or simply consists in, the painting having certain colors and shapes. That there should be this identity or constitution relation between the two is something that needs to be demonstrated and, so far, no reductive account along these lines appears entirely convincing. However, it turns out that when they are faced with substantive examples of this kind, many philosophers become immediately suspicious.\textsuperscript{144}

The dependence of final value shares the structure of evaluative supervenience but incorporates a thick notion of explanation or priority. The cost of going beyond evaluative supervenience in this way is that a certain mystery surrounding the dependence of final value will seem unavoidable, at least in so far as the ideas of metaphysical explanation or priority cannot themselves be analyzed. In other words, if we cannot give a proper account of these obscure notions, this will seem like an obstacle for any philosophical endeavor to better understand the concept of final value and the intuitions that govern its use. In the following subsection, we will consider some objections along these lines and spell out the extent to which they constitute a problem for investigations in the second-order domain. The objections are meant to establish that the concept of


\textsuperscript{144} Oliver is among them. He even calls for the concept of grounding to be banned after he states that “we know we are in the realm of murky metaphysics by the presence of the weasel words ‘in virtue of’” (1996, p. 48). I do not think that Oliver is necessarily wrong, but would argue that we are, in fact, in the realm of murky metaphysics whenever we engage in our evaluative talk and practices. To fully understand our intuitions about what it means for something to have final value, getting clearer about what those evaluative talk and practices commit us to is useful. To this end, I think that we had better keep the weasel words that Oliver warns us of.
Skepticism about Grounding

Hofweber (2009) objects to the notion of grounding and suggests that it is the result of esoteric metaphysics. This is metaphysical theorizing that proceeds from assumptions about a peculiar philosophical use of language. Hofweber argues that in so far as metaphysical grounding makes sense to begin with, it must involve a very special kind of explanation or priority that is not represented in our ordinary talk and practices. To put matters simply, he prefers an egalitarian metaphysics which proceed from more mundane understandings of what it means for one thing to depend upon another. Hofweber (2009, p. 268) also accuses proponents of grounding of relying on the tactic of bait and switch, whereby they point to some more mundane notion of explanation or priority and give an example of one thing depending upon another and then insist that we have a handle on grounding. To support his accusation, Hofweber goes through some of the substantive examples that are offered as illustrations of grounding. Disjunctions are true because at least one of their respective disjuncts is true, but this, he insists, is a straightforward logical implication. A couple is married because its members are married, but this, he insists, expresses a trivial conceptual truth. The density of an object depends upon its mass and volume, but this is similarly just a matter of entailment or analytic priority.

Now, the first two of these objections just mentioned appear problematic. Hofweber is right that the examples of disjunction and a married couple involve logical or conceptual equivalences, but this misses the point. As Raven (2012, pp. 691–692) stresses, if there is a logical or conceptual equivalence between two things, but we think that one of them explains or is prior to the other, then we are already on our way to accept something that shares the irreflexivity and asymmetry of grounding. This is precisely the case with the disjunction and couple example. That a disjunction is true and that at least one of its disjuncts is true are logically equivalent statements, but, intuitively, the latter explains or is prior to the former. That a romantic couple is married because its members are married is similarly a trivial conceptual truth, but, as Raven states (2012, pp. 691–693), it seems intuitively right to treat the former as being the case because the latter is the case. Neither of the two intuitions appealed to here are about supervenience, entailments, identity, or anything similar. We will concede the
point about density for now but will later see that this may not actually be such a bad example of grounding either.

Of course, Hofweber does not consider every example that has been offered in support of grounding but brings up a few to hint at a more general idea. The idea is that proponents of grounding need to provide examples of grounding that cannot be explained away by reference to more mundane dependence relations. Elements of this idea reoccur in the criticisms of Daly (2012), but he is much more explicit in his skepticism about the intelligibility of grounding. Daly (2012, pp. 91–96) mentions three common strategies for demonstrating the meaningfulness of grounding. The first strategy is to specify the formal properties of grounding, the second is to connect it to other understood notions, and the third is to give examples where it is thought to play a role. Regarding the first strategy, Daly notes that specifying the formal properties of grounding leaves its content undetermined and so it should not increase our confidence that we really understand its meaning. Regarding the second strategy, he suggests that the notions to which grounding can plausibly be connected either lie too close or too far away from it. When the notions are too close, he thinks that they are sufficiently like the concept of grounding to inherit its apparent obscurity. Among the examples mentioned by Daly are ‘fundamentality’, ‘metaphysical explanation’, ‘degrees of reality’, ‘mere aggregate’, and ‘existent’. He argues that when grounding is analyzed in such terms, what results are circular connections that fail to illuminate.

When the notions to which grounding is connected are too conceptually remote, on the other hand, they can be given more mundane interpretations, making the concept of grounding seem theoretically redundant. Examples mentioned by Daly include the notion of reduction. In the philosophy of science, when it comes to non-eliminative reduction, it is standard “to take a reduction of X’s to Y’s to be the discovery that X’s are identical Y’s” (2012, p. 93). Identity is of course symmetrical and reflexive, while grounding is not. “There cannot be pairs of facts,” Daly points out, “that stand both in the identity relation and in the grounding relation” (2012, p. 93). This shows that the notion of non-eliminative reduction cannot help us understand the notion of grounding, unless the former is given a less mundane interpretation that the

---

145 Daly’s strategy is a negative one. He questions whether the concept of grounding has been shown to be intelligible. See Daly (2012, pp. 83–88) for alternative strategies.
146 These notions are used by Schaffer (2009, p. 375) when he explicates the notion of grounding.
147 According to Daly (2012, pp. 93–94), the same problems crop up if we choose not to define grounding in other supposedly understood terms but rely on mere equivalences between grounding claims and other supposedly understood claims.
one Daly initially gives it. He speculates that there may be other understandings of non-eliminative reduction but suggests that it would be difficult to clarify them without resorting back to notions like ‘fundamentality’, ‘metaphysical explanation’, ‘degrees of reality’, ‘mere aggregate’, and ‘existent’. If this is right, then we are back to the problems mentioned above regarding the analysis of grounding. Lastly, Daly (2012, pp. 95–96) addresses the strategy of simply appealing to examples of grounding. He reiterates that such examples can only be indecisive, for those who do not understand grounding cannot be expected to understand the substantive examples that illustrate it:

Either [the grounding skeptic] finds that he does not understand the claims being made, and so the examples offered are as baffling as the general claim that some facts ground others. Or he finds that the examples are best construed as examples of relations of supervenience or identity, relations that are supposedly distinct from the relation of grounding. (Daly, 2012, p. 96)

Both Hofweber and Daly put emphasis on a certain class of uses of explanation and priority which they both regard as ordinary. They also argue that if the concept of grounding somehow goes beyond those uses, then it ends up being either theoretically suspect or will appear empty because it is incoherent. Before we continue, perhaps we should say something more about the more mundane relations that might be taken to speak against the concept of grounding here. Among all the possibilities mentioned in the literature, we find the relation of supervenience, entailment, identity, constitution, modal necessitation, truth-making, conceptual priority, realization, counter-factuality, determinables and determinates, types and tokens, functional roles and realizers, wholes and parts, sets and members, and sets and subsets. Some of these relations, as we have already seen, will not suffice to account for metaphysical explanation and priority. The reason is that they are not all asymmetric or irreflexive. However, even when we have removed relations like supervenience, entailment, or identity from the list, several other possibilities invite themselves. We shall see later that it is doubtful whether any of the others are any better at capturing the sense of

---

148 Koslicki (2015, pp. 311–312) is among those philosophers to also express the concerns that Daly expresses, in particular regarding the task of clarifying grounding by connecting it to other supposedly understood notions in its conceptual vicinity. She agrees that this is not very helpful and that those hoping to gain any clarity about notions such as priority and metaphysical explanation will mostly be left disappointed by the current literature.
priority which is involved in the dependence of final value. Before that, though, let us first consider a couple more objections to what has already been said.

Recall our friend who judges that a painting is good for its own sake and that its value depends upon its colors and shapes. In my view, the arguments for grounding are strongest when we consider cases such as these.¹⁴⁹ That our friend would only be committed to one of the more mundane relations when she judges that the painting is beautiful because of its colors and shapes is incredible.¹⁵⁰ Indeed, some of the more mundane relations seem to be of a sort that she is not committed to at all. The claim that by refusing to acknowledge her commitment in this regard, our friend has ventured into the realm of esoteric metaphysics, is equally strange given the ubiquity of this kind of normative explanation. It requires significant effort on the philosophical side of things to tease out the sense of dependence that is involved in judging one object to have final value because of this or that, but there is nevertheless nothing distinctly philosophical about that sense. Furthermore, we could also reject the skepticism shown by Hofweber with regard to esoteric metaphysics in the first place. As Raven points out, there are many accepted notions that we do not fully understand in our everyday lives: “Presumably, counterfactuals, Chomskyan syntax, autism, quantum states, and economic markets are all things we do not ordinarily understand. But we are to reject ground for that very reason. This double standard is irrational” (2012, p. 695). He continues:

Second, the association between being distinctively metaphysical and being something we do not ordinarily understand is weak. Some metaphysical relations (e.g. causation) are somewhat ordinarily understood. The initial examples of ground show that we have some ordinary understanding of them too. Lacking a strong understanding of a metaphysical relation, far from being a good reason to reject it, is a reason to illuminate it (ibid, p. 695).

With regard to the stubborn resistance of grounding to analysis, this should not convince us that we do not have some understanding of it. We do not know how to analyze knowledge, Raven (2012, p. 695) says, but this alone does not,

¹⁴⁹ Berker (forthcoming) shares my view in this regard. In his discussions, he convincingly argues for the need of grounding when it comes both to the first-order and the second-order domain. Many other interesting sources and relevant readings can be found in his paper.

¹⁵⁰ For some convincing arguments against specific reductions of grounding to a more mundane relation, such as truth-making and conceptual priority, see Raven (2012, pp. 690–670).
and should not, move us to conclude that there is no such thing as knowing. In addition, the preference shown by Hofweber for egalitarian projects would only make sense if there were clear examples of projects that were egalitarian but not also esoteric. Raven (2012, p. 696) is skeptical of the examples that Hofweber gives. Mathematics is not an egalitarian project, for it is inundated with complex notions that only have a passing resemblance to their ordinary counterparts. The notion of a set is a possible example, corresponding to a mundane and naïve notion of a collection. The mathematical notion of a set is not something definable in terms that are accessible to all and, according to Raven (2012, p. 696), it unavoidably calls for questions that belong to the more esoteric brand of metaphysics: Do mathematical statements that are undecidable in principle have determinate truth-values? This is a question about mathematical realism, which, although sensible, is presumably not part of egalitarian metaphysics.151

Our evaluative talk and practices involve, or commit us to, a metaphysical notion of explanation, corresponding to a thick notion of priority. Given that we use it, we seem to understand it, but perhaps this is an illusion. We should, at the very least, leave the door open for the possibility that common notions may be infected by metaphysical views that do not really make sense. Of course, even if this is true, it need not necessarily spell disaster for attempts to understand final value in terms of such notions. The attempts in question are meant to be describing the way that final value is represented in our evaluative talk and practices. We should leave the door open for the possibility that aspects of our evaluative talk and practices, such as they currently are, harbor incoherence. This does not mean that any analysis which appeals to the thick notion of explanation is incorrect. Daly at states that “only a term that is intelligible can genuinely contribute to our theorizing” and so “the utility of ‘grounding’ has to wait upon, and cannot prejudge, the verdict on whether ‘grounding’ is an intelligible term” (Daly, 2012, p. 96). It seems to me that Daly is here overstating his case. It would surely increase our understanding of our evaluative talk and practices e.g. if we learned that it harbors incoherence, exactly where that incoherence is to be placed, and why it arises to begin with.

In the face of the criticism by Hofweber and Daly, some philosophers might also want to insist that the concept of grounding is not called for by intuitions which take us beyond the more mundane understandings of explanation and priority, but rather that the concept is invited by those very understandings. The

---

151 See also Raven (2012, p. 697), for some responses to the more sardonic guilt-by-association arguments employed by Hofweber. I do not go into detail about these arguments here, for, quite honestly, I am not sure I understand them.
thought is that grounding picks out a broad metaphysical relation which underpins or is somehow abstracted from the more mundane understandings in question. Wilson (2014) and Koslicki (2015) independently offer arguments that, among other things, are meant to throw doubt on this line of reasoning. We should briefly consider the outline of these arguments, for the discussions they have inspired are very telling. Wilson and Koslicki suggest that there is no reason to think that there is a broad metaphysical relation that would, in any meaningful sense, subsume the more mundane understandings of metaphysical explanation and priority. These more mundane forms of dependence are simply not unified in a way that might justify grounding as a theoretical notion. At the very best, the concept of grounding ends up being little more than an umbrella term, standing as a sign for all relations that satisfy the formal earmarks that were highlighted in the previous section. Do we have some instance of metaphysical explanation or priority, involving a relation that imposes a partial order on its domain? If we do, then we are free to stipulate that we have encountered a case of grounding. However, Koslicki stresses:

My own view [...] is that classifying all of these phenomena as exhibiting grounding connections does not achieve much in the way of illumination. There are important and fairly obvious differences between these cases which have been obscured by creating the illusion that they are all connected via the single relation or operation of grounding. The important work of giving a positive account of the nature of the connections at issue still remains to be done, even after classifying all of these phenomena as exhibiting grounding connections; and we have not made much progress in that direction by applying a single label to what are evidently quite distinct phenomena (2015, p. 201).

If Koslicki is right about this, then it turns out that grounding, much like evaluative supervenience, may be too coarse-grained to be able to capture the dependence of final value. In fact, the dependence of final value may not be one thing, but several distinct phenomena, all of which need their own independent accounts. Wilson agrees with this general view of the heterogeneous nature of grounding. She says that “it is a myth that contemporary metaphysicians ignore metaphysical dependence and given the plethora of specific dependence relations on offer, there is not even a *prima facie* route from the failure of modal or diachronic relations to serve as grounding relations to the posit of any distinctive Grounding relation(s)” (2014, p. 539; my italics). According to Koslicki and
Wilson, the more mundane relations referred to earlier represent the kinds of tools that we need to use if we want a firmer grasp on how things end up depending upon other things. Otherwise, we have not been informative enough and simply skirted past the potential relevance of phenomena that we would have theoretical need for in any case.

I expressed incredulity regarding the suggestion that our friend would be committed to one of the more mundane relations when she judges that the painting is beautiful because of its colors and shapes. That our friend would venture into esoteric metaphysics by refusing to acknowledge her commitment in this regard seemed to me equally strange. The question is of course why the more mundane relations we have looked at will not do: What would be lost if our friend admitted that the explanation involves nothing more than one of the more mundane relations? Questions like this need to be answered if my personal incredulity is to have any weight. Again, even if we remove things like supervenience, entailment, or identity from the list of available mundane relations, there are many other candidates on offer which do satisfy asymmetry and irreflexivity, such as constitution, truth-making, conceptual priority, realization, counter-factuality, determinables and determinates, types and tokens, functional roles and realizers, wholes and parts. However, I would insist, along with philosophers like Cameron (2016), Schaffer (2016) and others, that even when they do impose a partial order on their domain, the more mundane relations fail to capture the relevant sense of explanation or priority. Curiously, Wilson herself seems to share this intuition:

So, for example, given that every x is a proper part of some Y, nothing follows about whether it is the parts or the wholes that are metaphysically dependent. The same is even more clearly true for other of the specific relations—in particular, type or token identity—that are frequently appealed to in investigations of metaphysical dependence. Moreover, as Fine suggested, ‘there is a real question, it seems to me, whether talk of more specific relations will be adequate to convey what we want to convey unless it is also coupled with a claim of ground’ (pers. comm.). On the ‘crucial appeal’ response, then, the proponent of Grounding maintains, first, that further (worldly) facts or assumptions must be in place in order for specific relations to serve as grounding relations, and, second, that these further facts/assumptions crucially involve an appeal to Grounding (2014, p. 558).
What is surprising about this is that Wilson is quite sympathetic to the objection but thinks she can get around it and capture the relevant sense of priority in terms of absolute fundamentality. Here the discussions start to take a rather strange turn. For Wilson suggests that we understand the thick notion of explanation and priority that concern us in terms of one of the more mundane relations, but with an understanding that we also assign a property of absolute fundamentality to one of their respective relata. For example, within the context of our discussions, the idea would be that we understand the relation between the final value of the painting and its colors and shapes in terms of one of the more mundane relations, such as material constitution, but add that the colors and shapes are fundamental. Proponents of grounding admit a mysterious and primitive notion into their conceptual scheme, but as Cameron (2016, p. 388) stresses, so does Wilson, since she agrees that such a notion is needed to make judgments about the structure of reality: “The main residual disagreement,” Schaffer summarizes, “is whether this primitive should be one of being-absolutely-fundamental, or the relative and linking notion of grounding (/being-relatively-more-fundamental-than-and-linked-to)” (2016, p. 157).

Cameron (2016. pp. 388–393) and Schaffer (2016, pp. 158–160) present some interesting objections to Wilson’s view, which we should briefly consider, since they help to further highlight the extent to which Wilson ends up being in the same boat as many of her opponents. Cameron (2016, pp. 389–393) argues that the notion of absolute fundamentality seems to be straightforwardly analyzable in terms of grounding. The idea is that one object is absolutely fundamental if, and only if, it is not grounded in any other object. If this analysis is on the right lines, it becomes difficult to see how Wilson’s account could really be that much more informative than the stories given by proponents of grounding. Cameron also points out that this account is not applicable to cases where there is meant to be structure all the way down and hence where no absolute fundamentality can be postulated. For similar reasons, her understanding is also not applicable to cases where there is meant to be structure between nonfundamentals—where one thing is more fundamental than another, although it has no absolute fundamentality. This becomes clear when we reconsider the example involving the painting, since the state of its displaying certain colors and shapes is almost certainly not a fundamental one. The point is that if everything that can be said in the preferred vocabulary of Wilson can be couched in terms of grounding, but the reverse is not the case, then her criticism of grounding is significantly weakened.

The objections raised by Schaffer are meant to show that Wilson does not heed her own theoretical advice, since her account of dependence fails to be as
informative or as unified she would like it to be. Our friend says that a painting is good for its own sake because of its colors and shapes. The colors and shapes metaphysically explain, and are therefore prior to, the final value which results from them. Wilson thinks that an understanding of this judgment in terms of grounding fails to be illuminating. Suppose, therefore, that we give a story following her formula. The state of the painting being good for its own sake is a constituent part of the state of the painting having certain colors and shapes. Moreover, the latter is postulated as being absolutely fundamental. Schaffer seems confounded by the claim that this kind of interpretation should be so much more informative than the previous one: “By my lights, both have said informative and potentially helpful things, though both should say more about how exactly the connection works, and they should do so by stating the connecting rule as precisely as possible” (2016, p. 160). Shaffer is right about this. The story resulting from an application of grounding and the story resulting from a combined application of a mundane relation and absolute fundamentality appear to me equally informative and amenable to expansion.

When our friend says that the value of the painting depends upon its colors and shapes, she seems to have ruled out many possibilities for alternative explanations (cf. Schaffer, 2016, pp. 147–148). We learn that the value of the painting is not completely unrelated to its colors and shapes. What is more, our friend has ruled out the possibility that the value of the painting depends solely on the context of its occurrence: The painting is not good for its own sake just because of the walls it hangs upon, the décor it happens to be a part of, its relative rarity in respect of other works of art, or the fact that it has been created by a historically significant artist. In spite of all this, Wilson insists that the explanation given is uninformative. This strikes Schaffer (2016, p. 148) as false and I agree. In fact, it seems to me obviously false. Wilson is of course right to suggest that our friend has not said everything she could have said, but this, Schaffer (2016, p. 148) also points out, would be no problem. Indeed, just like interpretations in terms of grounding may call for deeper stories about how and why the relevant dependence looks the way it does, so too can interpretations in terms of more mundane relations.

Schaffer (2016, pp. 159–160) stresses the latter point in his discussion. He argues that the combination of more mundane relations and the notion of absolute fundamentality on which Wilson relies would not mark an end to philosophical enquiry into matters of dependence.152 This point can be brought

---

152 Schaffer (2016, pp. 149–151, 162–167) also tries to give some hints as to what those deeper stories should, and might, look like. We need not go into more depth about those here.
out by reconsidering what would happen if we tried to adhere to the kinds of interpretation that Wilson prefers. Suppose we eschew any reliance upon grounding and interpret the explanation of our friend in terms of material constitution and absolute fundamentality. Clearly, the result would not be that everything that could be said about the relevant dependence had been said and that, in respect of this issue, philosophers can simply close up shop. Schaffer (2016, pp. 160–161) ends his critical remarks by noting that it is also unclear whether the notion of absolute fundamentality, or indeed the more mundane relations that Wilson talks about, really display more unity than grounding does.\footnote{See also Schaffer (2016, pp. 162–167) for logical and semantic work about grounding} It is not obvious that they do not also function as umbrella terms for more fine-grained relations.

On the whole, I think that arguments against the coherency or utility of grounding are relatively weak, and to the extent that they have merit, they need not provide any major difficulties for our reliance upon the concept in the chapters to follow. The following work is, for the most part, intended to represent an exercise in conceptual analysis, understood in a perhaps unconventionally broad sense, as the search for knowledge based on conceptual competence (Kipper, 2012, p. 9). To the extent that we will engage in metaphysical questions about what the world is like, we will do so only in an indirect manner. The questions we will ask will not be about whether final value exists or what final value is like, or indeed what grounding is like in any realistic sense. Instead, the questions we will ask could be understood as having the following form: If final value exists as it is represented by human beings, then what sort of thing could be said to exist and how could it be said to exist? Even if grounding turns out to be mysterious and less well understood than other concepts, it may still be of use when we look for answers to these questions. Whether grounding is a real relation, a part of the furniture of the world, is to my mind of little or no consequence in the current context.\footnote{It may still be argued that once we accept that the concept of grounding turns out to be redundant in respect of our best explanations of what the world is actually like, the motivation for finding out what we mean by the notion is also undermined. To this, I can think of no real response. The following work features an endeavor to clarify our everyday concept of final value, which, as we have seen, inescapably involves the concept of evaluative grounding. It does not feature an endeavor to show the reader why this clarification should be of any interest to her in the first place, although certain clues can be found in the introduction. If readers have made it this far, they presumably already find the topic of some interest, unless, perhaps, they}
Final value must depend upon other things, but it is unclear how this should be cashed out in closer detail. The argument has just been made that the claim that final value must depend upon other things is a conceptual one, but that the notion of explanation or priority that the claim involves is a metaphysical one. When we judge objects to be good or bad for their own sakes and that this is because the objects have certain other features, then we mean to say that the final value of the objects are explained by the features in question. Put differently, the features of the objects are prior to the final value whose instantiation they are responsible for. What this means was initially clarified by invoking the relation of evaluative supervenience. It was explained that the dependence of final value can be understood to mirror the structure of certain supervenience analyses, provided these were amended with the thick notion of explanation and priority which has just been mentioned. Otherwise the supervenience analyses ended up being too weak, as they were unable to capture the hyperintensional nature of the dependence of final value.

Once the attempts at analyzing the dependence of final value were left behind, we looked at some problems facing any attempts at understanding the relation in question. The first is that it is far from obvious why there should be a ban on mixed worlds. These were worlds where two objects differ only in terms of their final value. The second problem was that if we assume that this ban on mixed worlds is conceptual and that the dependence of final value is sufficiently strong, this seems to entail a reductionism whereby final value is identified with nonnormative properties. By way of responding to the first problem, it was reiterated that the ban on mixed worlds is conceptual and while there may be a story to tell as to why value concepts have the formal features that they have, the challenge of providing such a story is a perfectly general one that does not impugn the analyses of the dependence of final value which have been defended here. The second problem was avoided by considering inferences from necessary coextension to identity. The argument was made that although the dependence of final value renders the concept of final value necessarily coextensive with a disjunction listing its possible nonnormative basis, this does not necessarily entail that there is an identity between the two.
Finally, the last problem to be discussed concerned the concept of grounding, which was described as a more general version of the metaphysical concept of explanation or priority which has concerned us here. The dependence of final value is meant to incorporate a thick notion of explanation and priority, but several philosophers think that this notion is not so thick as to be interesting. In fact, they suggest that either the concept of grounding is devoid of content or otherwise theoretically suspect, on the one hand, or it should simply be understood in terms of more mundane relations, on the other hand. These more mundane relations include supervenience, entailment, identity, constitution, modal necessitation, truth-making, conceptual priority, realization, counter-factuality, determinables and determinates, types and tokens, functional roles and realizers, wholes and parts, sets and members, and sets and subsets. Here the argument was made that the concept of grounding does make sense and is not amenable to reduction to more mundane relations. What is more, even if grounding turns out to be incoherent, it may still be of use when it comes to getting a grip on our evaluative talk and practices. The idea is that these may well be incoherent as well.
Philosophers seemed convinced that if we just manage to figure out the nature of evaluative resultance, we would come to understand what it is for final value to depend upon other things (cf. Koslicki, 2015, p. 308). I noted in the introductory chapter that things are not quite so simple anymore. Besides the relation holding between final value and the factors that make objects good or bad for their own sakes, philosophers have become increasingly concerned with the other relations that are characteristic of complete normative explanations. More specifically, philosophers have begun to look very closely at the phenomenon of one factor contributing to other factors making objects good or bad for their own sakes. This widened perspective is partly due to the work of Jonathan Dancy (1993, 2004), who is among the philosophers chiefly responsible for shedding a light on the concept of background conditions. The background conditions of final value are the factors that contribute to its instantiation, without thereby being parts of what makes objects good or bad for their own sakes. The background conditions of reasons are the factors that contribute to their instantiation, without thereby being parts of the reasons themselves. The following chapter is concerned with the question what else background conditions could possibly be: How should we understand the role
that is played by certain factors when they are judged to contribute to normativity in this indirect sense?\textsuperscript{155}

It seems to me that when factors play the role of background conditions, they must have an indirect contribution to complete normative explanations that is either positive or negative.\textsuperscript{156} It is also reasonable to suppose, as Dancy (2004, pp. 38–42) does in his original discussions, that there are four distinct ways for background conditions to have such contributions: \textit{Enablers} contribute to other factors becoming effective value-makers or reasons, without themselves being among the value-makers or reasons; \textit{Disablers} contribute to other factors becoming ineffective value-makers or reasons, without themselves being among the value-makers or reasons; \textit{Intensifiers} contribute to other factors becoming more effective value-makers or reasons (in the sense that the resulting value or reasons become stronger than they would otherwise have been), without thereby being among the value-makers or reasons; \textit{Attenuators} contribute to other factors becoming less effective value-makers or reasons (in the sense that the resulting value or reasons become stronger than they would otherwise have been), without thereby being among the value-makers or reasons.\textsuperscript{157} The question posed at the end of the previous paragraph thereby offers a challenge to philosophers to formulate an account of background conditions that is capable of capturing all

\textsuperscript{155} Much gratitude is owed to Jakob Green Werkmäster for our many discussions (in private communication) about the phenomenon of background conditions and for helping me develop some of the ideas and objections found in this and the upcoming chapter.

\textsuperscript{156} When an object ends up having a neutral value and this is to be explained by the presence of a disabler, then I take that disabler to have had a negative contribution to the explanation.

\textsuperscript{157} We here assume that unless a factor is an effective value-maker it is not a value-maker at all. The concept of making something good or bad is in this sense treated as a success concept. However, we should be careful not to assume that intensifiers always have a positive effect on final value in the sense that they help make things finally better and that attenuators always have a negative effect on final value in the sense that they help make things finally worse. What intensifiers do is strengthen whatever sort of value happens to be instantiated, while attenuators have the opposite effect. This means that intensifiers may serve to strengthen the final badness of an object, just as attenuators may serve to weaken the final goodness of an object (cf. Nozick, 1968, pp. 36–37). It is interesting that when the distinction is seen from this perspective, disabling starts to look like a radical form of attenuation. On one suggestion, a factor is a disabler if it attenuates the effectiveness of a value-maker to zero (it disables it from being an effective value-maker). The object can still be good or bad for its own sake, but not just in virtue of the disabled value-maker. On this suggestion, a factor is an enabler if it intensifies the effectiveness of a value-maker from zero to some positive value. I will make no assumptions regarding the reducibility of enablers and disablers to intensifiers and attenuators in what follows. However, I will assume that all the phenomena just mentioned belong to the same general category and so cry out for a common account.
the phenomena just mentioned. The following chapter features an attempt on my part to meet this challenge.

I will develop an account of background conditions that is supported by the assumptions mentioned and defended in the introduction. In addition to trying to capture the phenomenon mentioned above, the account in question has two theoretical aims: The first is to be as accommodating as possible when it comes to the substantive intuitions that surround the phenomenon of background conditions. My reason for wanting to pursue this aim is that I share many of the substantive intuitions in question. To me there does seem to be something reasonable about the general view that the factors that help contribute to the applicability or accuracy of explanations need not themselves be parts of the explanations (cf. Dancy, 2004, p. 45). This general view seems to me sufficient to at least reinvite the concept of background conditions into our evaluative talk and practices. The second aim is to avoid the conclusion that background conditions must involve their own special types of relations, which sit alongside evaluative resultance. The reason why we should want to pursue this aim is that evaluative resultance is mysterious enough. If it can be avoided, furnishing the dependence of final value with additional *sui generis* relations would be a bad idea. To these ends, the chapter is divided into four parts, the first of which consists in a brief preamble that is meant to clear the ground of any irrelevant senses of background conditions, thereby introducing the phenomenon of interest in a general and noncommittal way.

Following the preamble is a discussion focusing on the deflationary view which suggests that background conditions should not be taken seriously as a category of factors that are distinct from value-makers or reasons. The first version of the view takes the form of a *pragmatic interpretation*, which understands background conditions as mere conversational tools that are meant to put emphasis on what we deem especially interesting in complete normative explanations. Where this emphasis is most naturally put depends upon our interests and expectations, as well as the communicative norms that they help shape. The second version of the view that we will consider is an *overriding interpretation*, which understands background conditions in terms of the concept of an exclusionary reason. Roughly put, an exclusionary reason is a reason to treat other facts as if they were not reasons, or to downplay them for the purposes of normative deliberation (Raz, 1990, pp. 37–40). We shall see that the pragmatic interpretation and the overriding interpretation can be combined to give an attractive picture of why we come to think of certain factors as background conditions. However, the argument will be made that to complete the picture of what it then means to think of factors as background
conditions, we may also need to look elsewhere. One of the biggest problems is that the deflationary view fails to accommodate some of the most central intuitions that underpin the phenomenon of background conditions.

After discussing the deflationary view and its problems I will begin to formulate and defend my preferred account of background conditions. On this account, the background conditions of final value are the factors in virtue of which, or as a result of which, facts become reasons for attitudes. Whenever facts are afforded the normative property of being a reason, there must be some other factor present that makes the facts into reasons. These factors are what should be regarded as the background conditions for the resulting final value. This account will not only deepen our understanding of the distinction between background conditions and value-makers, but of the nature and behavior of reasons as well. Its fruitfulness will be illustrated by showing how it may be applied to illuminate the concept of derivative value, which is often invoked in moral philosophy but rarely explicated. The idea is that the concept of derivative value should be understood in terms of reasons for attitudes that result from the presence of another reason. In other words, when the instantiation of the normative property of being a reason in one fact helps explain why another fact is also afforded with that same property, then what results might be a derivative value. The wider implications that this understanding has are briefly highlighted with additional examples taken from the first-order domain.

Finally, it should be mentioned that since I make use of the fitting-attitudes analysis of final value, I will often permit myself in what follows to glide from talking about final value and talking about reasons for attitudes. I see this as relatively unproblematic. Among the substantive intuitions that I wish to accommodate is the one which states that the background conditions of final value are something very different from value-makers. Because I accept the fitting-attitudes analysis of value, there is another equivalent way of couching the substantive intuition in question: The background conditions of final value are something very different from the facts that constitute our reasons for attitudes. Another substantive intuition that I wish to accommodate is the one which states that the background conditions of reasons themselves are also something very different from the facts that constitute the reasons in question, whether they are reasons for attitudes or not. Although the main purpose of this chapter is to understand the background conditions of final value, then, what is offered is a unified account that applies just as much to reasons.
Preparing the Ground

Even though philosophers have become increasingly concerned with how factors can contribute to complete normative explanations without being value-makers or reasons, it has been noted that very few have ever attempted to offer a positive account of background conditions (Bader, 2016, p. 28). Even Dancy, who is the first philosopher to put a lot of emphasis on the concept, never makes a detailed attempt to explicate it. Instead, he starts of his seminal discussions of background conditions somewhat acquiescently, by stating: “If one cannot explicate a philosophically significant concept, there may be other ways to give people a sense that the concept is itself in good order and that they have a reasonably clear grasp on it” (2004, p. 38). One such way involves Dancy trying to give us a sense of the concept he has in mind is by using substantive examples, two of which will play a significant role in the investigations to follow. The substantive intuitions to which the examples appeal are precisely those intuitions that my preferred account is ultimately meant to accommodate.\(^{158}\) The first example focuses on value while the other focuses on reasons:

---

**The Joke:** An amusing joke is told in the presence and at the expense of a work colleague. The joke is made amusing only by the fact that the joke is witty and sarcastic. However, had the colleague not been there to hear it, the joke would not be amusing at all. It would then only have been cruel and inconsiderate. This suggests that the presence of the colleague enables the fact that the joke is witty and sarcastic to make the joke amusing.\(^{159}\)

---

\(^{158}\) For a third example which we shall not focus on, not because of its implausibility as much as because of its complexity, see Dancy (2003, pp. 632–633). The example revolves around the suggestion that close relationships would have no final value if they were merely a sham, yet the fact that a specific instance of, say, friendship is not a sham is not part of what makes the friendship good for its own sake. To flesh out the example, Dancy makes wonderful use of examples from the world of fiction by focusing on the film *The Truman Show*. The main protagonist of the film is a man who has grown up in front of cameras in an enormous film studio, completely unaware that everyone around him is an actor playing a role. Even his mother, father, wife, and childhood friend are pretending to be someone that they are not. The close relationships of the main protagonist, we are to assume, do not have the significance that they otherwise would have, if they were not just the result of a script.

\(^{159}\) It is important to remember that the property of being amusing is here understood as an evaluative one. For a joke to be amusing is not the same thing as its tending to elicit laughter or amusement, then, but is rather the same thing as there being reasons to laugh or be amused by the joke. The example is originally borrowed from the discussions of Dancy (2004, p. 172).
The Promise: One person promises another person that he will attend her birthday party. The fact that the promise has been made constitutes a reason for him to attend her birthday party. However, if he was completely unable even in principle to attend the birthday party, the fact would no longer constitute a reason for him to do so. This suggests that his ability to fulfill his promise enables the fact that he has made it to be a reason.160

The intention here is to tease out the intuition that even though value-makers and background have a similar logical and counterfactual relationship to the presence of final value in these examples, background conditions have a less direct effect on final value than value-makers do.161 This is why it would be misleading, for example, to say of something like an enabler that it enables objects to have final value, since what it does is really to enable other factors to become effective value-makers.162 A somewhat related observation that might be gleaned from the joke example is that value-makers play a more essential role in normative explanations than background conditions do. The suggestion is that when one factor makes an object good or bad for its own sake, then it does not necessarily follow that there are background conditions present that explain why this is so. This would indeed make background conditions very different from

Dancy has another version of this example where the joke is not amusing unless the person it is about is not present to hear it being told, but he does not think that this is as illustrative of the concept he has in mind: “In these cases, I would think, their presence is part of the ground for the cruelty; it is part of what makes telling the joke cruel rather than delightful. So here we are not dealing with enablers at all. But there are other jokes which are only amusing if their butt is present, and which, in that case, are amusing to everyone present, to the butt as much as to everyone else; if the butt was not there, that would take all the fun away” (2004, p. 172).

This example is also borrowed from Dancy (2004, pp. 38–39). A similar example is used by Nozick (1968, p. 36) in his discussions about background conditions. He imagines that one person makes a promise to another, but that the latter then releases the former from her promise. According to Nozick, the feature of the situation that the person does not keep her promise has then been canceled from a normative standpoint. The feature would no longer make it bad that the promise is not kept—assuming that no third party is harmed as a result.

The substantive examples that have just been cited could be modified to illustrate the concept of attenuating or intensifying instead. For instance, there may be a joke such that it becomes either more amusing or less amusing when the person it is about is present as it is being told.

Just like value-makers, background conditions are typically understood as being properties. The reason I keep talking about them in very generic terms as being factors, is that I have yet to see a reason why I should not be a pluralist and assume that background conditions can belong to many ontological types. They can be properties, objects, relations, structures, tokens of states of affairs, facts, and so on.
value-makers, because, as we have seen, objects cannot have final value without there being something present that makes them good or bad for their own sakes. Now, the view that background conditions do not have to be present even in complete normative explanations might seem incompatible with my preferred account, since, on that account, background conditions are the factors that make facts reasons for attitudes. I have also admitted that whenever a fact is a reason there must be something that provides it with a normative status. However, I shall later argue that background conditions are only those specific factors that afford facts their normative status without being parts of the facts themselves. This will be shown to entail that my preferred account is in fact compatible with the view that background conditions are not as essential to normative explanations as value-makers are. Another observation in the vicinity of this topic regards the possibility mentioned by Dancy (2004, p. 91) that there are background conditions for background conditions.

The general view that the factors that help to make an explanation correct, need not themselves be part of the explanation, leaves the door open not only for things like enablers, but also for the enablers of enablers. Whereas enablers are meant to contribute to other factors becoming effective value-makers, enablers of enablers contribute to other factors becoming the enablers of effective value-makers. For example, there might be a factor that explains why it is of relevance to the joke that the colleague is present in the first place. It is admittedly unclear what this something could be, but here is one suggestion: The joke is amusing in virtue of the fact that it is witty and sarcastic, but the joke would not be amusing if the colleague was not present to hear it. However, the presence of the colleague would not be required if the joke was not told in a particularly sneering tone. That the joke is told in a particularly sneering tone, then, is neither something that makes the joke amusing, nor something that enables the joke to be so. If this example is on the right lines, then it also seems to speak in favor of the observation that background conditions often play a less direct role in normative explanations than value-makers do.

Before we move on we need to distinguish the sense of background conditions that will be relevant to the following discussions from those that will

---

163 The argument will also be made that even if background conditions are not as essential to normative explanations as value-makers are, this should not be taken to mean that background conditions do not play an essential role within the first-order domain.

164 Intensifiers and attenuators are a possible exception to this intuition, since it seems to make less sense to say that one factor could make another into a more or less effective enabler or disabler. On the other hand, it does seem to make sense to say that one factor could make another into a more or less effective intensifier or attenuator.
not. Philosophers of language might look at the promise example and draw associations with the concept of felicity conditions, as utilized in J. L. Austin’s (1962) formulation of speech-act theory. To utter a promise is to make something the case, rather than describe any state of affairs. Such an utterance can therefore not be true or false in a strict sense, but it may nonetheless be characterized by certain felicity conditions. For example, if it turns out that we were in principle unable to fulfill the promise that we made, then this could be seen as an indication that we failed to make the promise in the first place. It would be like the pronouncement of two people as married outside of the appropriate context of a wedding ceremony and in the presence of a licensed solemnizer. The distinction between having an ability to fulfill a promise serve as a felicity condition for the promise and having it serve as a background condition for a reason to fulfill the promise is a difficult one, but that there is such a distinction seems rather obvious. Something could be a felicity condition without being a background condition and vice versa.

There are some additional understandings of what background conditions are that I would like to put to the side. Let us first imagine that we are standing in a dark room and are unable to appreciate the intricate piece of art that is hanging on the wall right in front of us. Suddenly, someone decides to turn on the lights and the vibrant colors and joyful patterns of the painting jump out at us. We can now see just how beautiful the painting really is. The question to be considered is whether we should regard the presence of light as an enabler for the value of the painting in this specific case. It may be difficult to see why not. Ultimately, it is of course a matter for first-order theories to determine whether any specific factor plays the role of an enabler (cf. Dancy, 2004, pp. 51–52). That said, it needs to be kept in mind that background conditions are meant to exert their influence on value-makers themselves and not just on our ability to perceive or appreciate the value-makers. It seems reasonable to insist, then, that the light does not constitute the kind of background condition that we will hereafter be interested in, if all it does is make it possible for us to see how beautiful the painting was all along.

Imagine next that we have a beautiful flower standing on our windowsill that is exposed to too much sunlight. The result is that the flower starts to wither and become brown, thus losing its previous value. The question to be considered is whether the right amount of sunlight could be regarded as an enabler for the value of the flower in this case. It is not as easy to dismiss this example as the previous one, because the right amount of sunlight would, by causal interaction, ensure that the flower retained the factors that make it beautiful. There is therefore a sense in which the proposed enabler does have a direct effect on the
relevant value-makers and not just on our ability to perceive or appreciate them. The problem here is of course that this does not appear to be the right type of effect. To see this, we need only think back to the joke example and make some helpful analogies. The presence of the colleague does not have the same effect on the joke as too much sunlight has on the properties of the flower. The difference between the kinds of background conditions we shall hereafter be interested in and those that are involved in this modified version of the flower example is that the latter involves a causal phenomenon.165

The kind of background conditions that we will be interested in, then, are not the factors that merely make things the possible objects of our awareness or appreciation (as in the case of the painting being lit up by lights), nor are they the factors that causally influence objects in the sense that they enable them to have value-makers in the first place (as in the case of the flower being exposed to too much sunlight). Dancy recognizes the latter point when he says that we are looking for “examples where what is enabled is not the presence of the valuable feature, but the value of a present feature” (2004, p. 172). He also seems to recognize something very similar to the former distinction when he says that “what is required for one to know on the basis of evidence need not itself count as more evidence” (2004, p. 49). What Dancy is referring to in this quote is presumably background knowledge and skills that enable us to learn something new, without themselves constituting a reason to believe in anything new. I take it that this is just a special case of the more general epistemic phenomenon that has just been mentioned. To make things simple moving forward, we might continue to refer to the background conditions that we will hereafter be interested in as ‘background conditions’, while the other two could be referred to simply as ‘epistemic background conditions’ and ‘contributory background conditions’.

The Deflationary View

The deflationary view insists that background conditions should not be taken seriously as a category of factors distinct from value-makers or reasons. The intuitions appealed to in the kinds of examples we have been looking at should

165 Causal relations are typically taken to hold only between events and states of affairs that are at least partly nonnormative. It may be difficult to make clear exactly what this means, but intuitively, at least, it seems as if the fact that a property is a value-maker does not qualify.
therefore be explained away rather than captured. This is presumably seen as necessary to attain theoretical economy and metaphysical parsimony. It turns out that there are several interesting ways in which the deflationary view can be spelled out. To start with, I think that it would be helpful to note of some of the similarities between background conditions within the normative domain and background conditions as they are used in other areas of philosophy. There are several reasons for doing this, among which is the fact that discussions of background conditions within those other areas of philosophy have tended to be more developed and sophisticated in their details. Another reason for doing this is to hint at the sort of work that needs to be done to make the deflationary view look at all plausible. We will more specifically be borrowing from discussions in the philosophy of science, where the relevant distinction is formulated in terms of background conditions and causes:

*The Match:* A match is dragged along a rough surface and the friction ignites the chemical elements at the tip of the match. The result is that the match catches fire. The match would not have caught fire unless there was oxygen in the air and yet the presence of oxygen is not among the relevant causes. Instead, the presence of oxygen in the air enables other factors to play the role of causes for the match catching fire.166

*The Plane:* A plane suffers from an unfortunate series of mechanical failures, made even worse by certain pilot errors. The result is that the plane crashes into a lake. The plane would not have crashed into a lake unless it was subject to gravitational forces and yet gravity is not among the relevant causes. Instead, the plane being subject to gravitational forces enables other factors to play the role of causes for the plane crashing.167

The general intuition is that the factors that make an explanation correct or fitting need not themselves be part of the explanation.168 If the cases just described are taken at face value, this intuition does not just seem to hold true for normative explanations, but for causal ones as well. There may be factors

166 The example originally comes from Hart & Honoré (1985/1959, p. 35).
167 The example is borrowed from Cheng & Novick (1991, p. 84).
168 Some sources suggest that the distinction between causes and their background conditions, like so many other things in philosophy, traces back to the dialogues of Plato (Dancy, 2004, p. 45).
that explain how other factors can become causes, then, without the former factors being causes themselves. Interestingly, although the two examples we have just looked at may seem intuitively plausible, the prevailing tendency within the philosophy of science is to regard such cases with skepticism. The consensus among philosophers seems to be that the distinction between causes and background conditions is a superficial fiction that lacks any firm basis in objective reality.\footnote{This way of looking at the distinction between background conditions and causes has a surprisingly long tradition. John Stuart Mill famously states that distinction can be shown to lack any objective grounds by “the capricious manner in which we select from among the conditions that we choose to denominate the cause” (1973/1846, vol. 7, p. 329). Regardless of how many background conditions we may identify, “there is hardly one of them which may not, according to the purpose our immediate discourse, obtain that nominal pre-eminence” (1973/1846, vol. 7, p. 329). The soon to be discussed pragmatic interpretation of the deflationary is meant to add to this picture and show that, while it may not be grounded in objective reality, the distinction between background conditions and causes is not quite as capricious as Mill once feared. The distinction exhibits a certain predictable pattern that can be modeled by way of normalcy conditions and conversational implicatures.} Despite this, attempts have been made to show that there is an interesting and predictable pattern to its application. This has been done by identifying, explaining, and to some extent justifying the many ways in which the distinction tends to figure in ordinary causal explanations. The result of these attempts has been the development of a broad interpretation of the deflationary view which takes the distinction between causes and background conditions to mainly be a pragmatic instrument. We shall consider the pragmatic interpretation of background conditions below and will see how it is meant to apply in both the causal and the normative case.

On the Pragmatic Interpretation

The concept of background conditions is seen by some philosophers as a tool of emphasis, which can help speakers to shed a light on what is deemed to be salient in ordinary causal explanations. Where this light is most naturally directed has to do with the interests and expectations of speakers and their hearers, but not with what can plausibly be regarded as causes in the first place. Accordingly, the presence of oxygen can be found among the causes of the match caught fire, just as gravity can be found among the causes of the plane crash. The reason that these things may be omitted from ordinary causal explanations has to do with what people tend to care about. The pragmatic interpretation thereby suggests that background conditions are constituted by
factors that have relevance for the causal phenomenon of one thing explaining another, even if they may lack, for instance, practical importance with respect to the human activity of giving causal explanations.\textsuperscript{170} Of course, if it is true that the patterns in accordance with which we use the distinction between background conditions and causes are shaped by interests and expectations in this way, there should be certain predictions that we can make about how the distinction tends to be used.

Among the predictions that we should be able to make is that people will place their emphases quite differently when met with modified versions of the examples. The idea is that expectations and interests are sensitive to context in such a way that how people give causal explanations should be subject to controlled manipulation. Indeed, it seems very plausible that this would happen regarding the first example involving the match. It has been pointed out that if the match catches fire in a setting which is mistakenly assumed to be free of any oxygen, then things will appear to us very different. We would then not be surprised to find that people start placing their emphasis differently and might include the presence of oxygen in their causal explanations (Hart & Honoré, 1985/1959, p. 35; Cheng & Novick, 1991, p. 93). Whether this line of reasoning would seem as plausible in the plane example is less clear. More effort is certainly required in order to imagine a realistic example where the plane crashes in an environment where gravity was assumed to be a less significant factor than usual. If such a case could be imagined, on the other hand, then perhaps we would not be surprised to find that people start placing their emphases differently by including gravity among the relevant causes.\textsuperscript{171}

Aside from giving us predictions that we should be able to test empirically, the pragmatic interpretation of the deflationary view has numerous other

\textsuperscript{170} I assume that when the National Transportation Board writes reports about plane crashes, the fact that aircraft are generally subject to gravitational forces is usually not mentioned.

\textsuperscript{171} For some empirical studies along these lines, see the discussions of Cheng & Novick (1991, pp. 101–117). The following remarks about normalcy conditions and conversational considerations are inspired by, and at several points follow, the general overview they provide. However, we shall not be considering their detailed criticism of the normalcy conditions or conversational considerations, unless said criticism should prove relevant in the case of value. We shall also forego consideration of their alternative view of background conditions, because this view involves a probabilistic and predictive component. This component seems fitting in the causal case, but less so in the normative case. There is a tradition in the philosophy of science to understand causes in probabilistic and predictive terms, but this is obviously not applicable to the relation between makes and value. The latter relation is supposed to be stronger and is not generally ascribed so that we may make predictions about future occurrences. If the view of background conditions defended by Cheng & Novick is correct, then, this reveals a major point of disanalogy between the causal case and the normative case.
benefits. Among them is the fact that there are many ways in which the interpretation can be cashed out in an even greater theoretical detail. Hart & Honoré (1985/1959) were among the first to note that interests and expectations tend to depend upon what is already known and what would be considered commonplace in certain circumstances. The presence of oxygen might not be mentioned in the first version of the match example, because it is a common part of circumstances in which we are asked to explain why things like matches catch fire. We would be less surprised to see it mentioned in the second version of the example, because the presence of oxygen is not a common part of circumstances involving carefully contained laboratory settings. Gravity is not mentioned in the first version of the crash example, because it is a common part of circumstances where we are asked to explain why anything falls from the sky on this planet. Of course, this means that it might be mentioned in other versions of the example, if there is a more exotic kind of circumstances where gravity is not as expected.

Cheng & Novick (1991, pp. 85–87) suggest that normalcy conditions of this kind can be either statistical or idealistic. The statistical normalcy conditions predict that background conditions are the factors that people would consider to be normal parts of an occurrence. The idealistic version of normalcy conditions, on the other hand, has it that background conditions are the factors that do not accord with certain norms. Kahneman and Miller expound upon the latter in a pedagogical way when they write that: “Peculiar behaviors of cars observed on the road are frequently ‘explained’ by reference to the drivers being young, elderly, or female, although these are hardly unusual cases. The default value for an automobile driver appears to be a middle-aged male, and driving behavior is rarely explained by it” (1986, p. 149). We need not go into further detail about the nature of normalcy conditions here. The crucial point is that this concept can help us to accommodate the intuition that when people ask for causal explanations, it is because something unexpected has happened and they are curious about what sets an occurrence apart from others (Cheng & Novick, 1991, p. 85). The extent to which this is also true within the normative domain will be considered soon.

Other philosophers have pointed out that the interpretation under consideration can be made more nuanced by incorporating insights about the

pragmatics of language (Cheng & Novick, 1991, pp. 88–91).173 Among the more relevant insights are those offered by H. P. Grice (1975) concerning conversational implicatures. To distinguish between distinct aspects of meaning, Grice formulates a series of rules that are implicitly accepted by the contributors to meaningful conversations.174 Two such rules state that contributions to meaningful conversations should only be made as informative and relevant as is required by context.175 The general idea, then, is that treating a factor as a background condition might be a method of conforming to the rules in question (Cheng & Novick, 1991, p. 89). This would fit very nicely with what has just been said about normalcy conditions in the previous paragraphs, since what people deem informative or relevant is presumably affected by their interests and expectations (Swanson, 2010, p. 226).176

Much more could probably be said about the theoretical details involved in the pragmatic interpretation. However, the brief introduction that I have just given should suffice to convey a sense of the main theoretical paths that must be explored to make the pragmatic interpretation seem plausible and for it to constitute a proper defense of the deflationary view. Once amended with the pragmatic interpretation, the deflationary view provides us with a very interesting theory, rich in philosophy as well as empirical content. On this theory, the distinction between background conditions and causes is a mere pragmatic tool, used in causal explanations to put emphasis on what speakers deem salient. Rather than being determined by mere caprice, this emphasis shifts in accordance with our interests and expectations, in a predictable pattern that can be modelled by way of things like normalcy conditions and conversational implicatures. The question we should now turn to regards the extent to which this style of approach is as applicable to the normative case as it appears to be in


174 Grice regarded all conversational maxims as subsumed under a general principle: “Make your conversational contribution what is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (1975, p. 45).

175 Swanson (2010) expresses this idea by stating that although background conditions are themselves parts of causal chains, our interests and expectations result in background conditions not being viewed as good representatives of those causal chains. He writes that good representatives are those that “provide conversational participants with enough information about the represented causal paths for the purposes at hand” (ibid, p. 226).

176 It may seem obvious if not trivial that what contributions are deemed informative or relevant is affected by what we consider to be part of the commonplace. However, Cheng & Novick (1991, pp. 89–90) are very careful to stress that we should not assume that normalcy conditions and conversational considerations always overlap, for they apparently do not.
Although I have my doubts, there is certainly no shortage of philosophers that have argued in favor of a broadly pragmatic interpretation within the normative domain. Among those philosophers is Caj Strandberg, who describes his view of the background conditions of value as follows:

Strandberg thinks that the factors that speakers tend to think of as background conditions for value are the factors that play a role in making things good or bad for their own sakes (either by being makers themselves or by being parts of more complex makers), but since they play a less direct role than certain other things, they tend to be completely left out of our normative explanations. It is unfortunately unclear how he understands this less direct role that is played by background conditions, but one possibility is that background conditions are those makers that commonly occur in resultance bases, but that always require the presence of other things to be makers. Some factors are independently value-making in the sense that they do not require the presence of additional properties to give rise to final value. Other factors are not independently value-making in this sense, but only gain a direct influence on final value once they occur in a more complex maker. In a later work Strandberg states more explicitly that this is how he understands the background conditions of value. He says that even if Dancy is right that “a non-moral property which he classifies as an enabler cannot make objects have moral properties”, this does not entail that “a set of non-moral properties which includes such a non-moral property cannot have this function either” (2008, p. 142). This adds a substantive element to the

177 Among those to have argued that a broadly pragmatic interpretation is applicable in the case of value or reasons, we find Raz (2006, pp. 110–113), McKeever & Ridge (2006, pp. 71–75), and Rønnow-Rasmussen (2011, p. 150).
approach under consideration that is very interesting, but the approach remains
a pragmatic one in the main. Another more recent example of a philosopher
who takes this general pragmatic approach to background conditions within the
normative domain is Fogal, who describes his view of the background
conditions of reasons as follows:

Call any cluster of facts that taken together—and only taken
together—are normatively relevant (in some particular way, to
some degree) to some action or attitude a "normative cluster." For
reasons having to do with communicative efficiency as well as our
informational limitations, we rarely mention entire normative
clusters. Instead, we single out one or two particularly relevant
and/or accessible facts and ascribe responsibility to them. In
essence, the cited facts function communicatively as
representatives of, or proxies for, the normative clusters to which
they belong, or are otherwise related (2016, p. 91).

This more technical explanation of background conditions comes even closer to
mimicking the pragmatic interpretation as it is applied to the causal case
above. In the quoted paragraph, Fogal explicitly endorses the idea that what
speakers think of as background conditions for reasons are those factors that we
leave out of emphasis for broadly communicative purposes. In short, they are the
things that speakers are not very interested in hearing about and which are for
this reason left out of their normative explanations. For all that, though,
applying the pragmatic interpretation within the normative domain in this way
is not as straightforward as this line of reasoning may lead us to think. To begin
with, it seems obvious that to work properly, the interpretation needs to be
adjusted and amended, lest it end up assigning a normative role to the wrong
sort of factors. That it assigns a normative role to the wrong factors would result
in the pragmatic interpretation forcing us to make some awkward decisions,
which, intuitively, it seems we should not have to make. These problems will
come into focus in the next section, where we will begin to take a closer look at
the overriding interpretation, which offers additional explanatory tools to be
used by the deflationary view. Before that, though, let us recapitulate and state
once more what the pragmatic interpretation commits itself to, for it is quite
important that we know what is at stake.

178 Indeed, Fogal (2016, p. 88) admits to being inspired by, and borrowing from the views of, the
previously mentioned Swanson.
In the joke example, the presence of the colleague is claimed to be part of the factors that make the joke amusing. Because we are assuming that the fitting-attitudes analysis is correct, this is equivalent to saying that the fact that the colleague is present is part of that larger fact which constitutes our reasons to be amused by the joke (recall that this larger fact also incorporates the fact that the joke is witty and sarcastic). Any substantive intuitions we may have suggesting that we should not treat the fact that the colleague is present as a value-maker, or equivalently, as part of the reason we have to be amused by the joke, is to be explained away rather than accommodated. Similarly, in the promise example, the fact that one person is able to fulfill his promise to his friend is seen as part of that larger fact which recommends to him that he keep the promise (recall that this larger fact also incorporates the fact that the promise has been made to begin with). Again, any substantive intuitions we may have suggesting that the fact that the person is able to fulfill his promise is part of the relevant reason should be interpreted through the lens of our interests and expectations, and the communicative norms that these have helped shape. Let us now move on to some of the shortcomings of the pragmatic interpretation and consider some of the ways in which the deflationary view can be further amended.

On the Overriding Interpretation

We saw in the introduction to this work that final value can be both intrinsic and extrinsic relative to its resultance base. Despite this possibility we should be free to adopt a first-order view which states that all final value is in fact always either intrinsic or extrinsic. One of the problems for the pragmatic interpretation is that it can potentially come into conflict with a first-order view of this kind. For the sake of simplicity let us focus on the possibility that all final value is intrinsic. The pragmatic interpretation would then seem to present us with the following dilemma: Either the first-order view that final value is always intrinsic is wrong or the first-order view that background conditions can ever be extrinsic is wrong. The reason is that, on the pragmatic interpretation, background conditions are themselves value-makers. This means that if a final value is intrinsic relative to its resultance base, then the relevant background conditions for that final value must be intrinsic as well. My first objection toward the pragmatic interpretation, then, is that this dilemma does not appear to be a choice that we should have before us. More modestly put, there are reasons to prefer any account of background conditions that avoids putting us in this kind of theoretical choice situation.
A closely related problem with the pragmatic interpretation concerns the issue of explanatory complexity. The idea is that once we include all the background conditions for a given value among its makers, what results might be a resultance base of enormous complexity. This is worrying given the fact that such a resultance base could perhaps not be represented in the right way in the intentional content of our attitudes. This means that background conditions are perhaps not just selected based on our shared interests and expectations but based on what can be represented in speech and thought. One way of responding to this worry would be to distinguish between the actual and merely conceivable explanatory complexity that might result from the pragmatic interpretation. We can imagine that the resultance base for a given value turns out to be enormously complex, but that does not mean that it must be. We can always be more restrictive when we look at these things from the perspective of the first-order domain. The problem with this response is that for any given value many of us could, after enough careful reflection, identify a great deal of things we would consider to be background conditions. Once we are told that background conditions are themselves just value-makers, we would apparently be forced to either change our minds completely about the relevance of those factors or expand the relevant resultance base. Once again, the point I wish to make here is that there are reasons to prefer any account of background conditions that avoids putting such choice before us.\(^{179}\)

There is also the worry that background conditions are not always left out of emphasis merely because we are not interested in hearing about them, supposedly despite the relevance they really have. Instead, some background conditions are left out of emphasis precisely because they do not have any direct relevance to begin with, although they might play a broader role in affecting value-makers. This intuition can be given some measure of support by reconsidering one of the examples with which this discussion started. It might seem reasonable that it should be an expected and normal part of circumstances where promises are made that the person making the promise is also able to fulfill them. However, is it as reasonable to say that it would be an expected and normal part of circumstances where a joke is told at the expense of a colleague that she is also present to hear it? If someone was informed of the role played by

\(^{179}\) One might try to soften the blow of the dilemma by appealing to ideas about indirect value-making. The argument could be made that the value-makers we ordinarily treat as background conditions are features that only become value-making when they occur in conjunction with other features. Gratitude is owed to Toni Rønnow-Rasmussen and Włodek Rabinowicz for pressing this point (in private communication).
the presence of the colleague, would she react with the same disinterest as someone who was told about the role played by oxygen for the lighting of matches? If someone insisted on including the presence of the colleague in her explanation why the joke is amusing, is her only mistake that she has violated a communicative rule of informativity? My preferred answer, which is driven by the intuitions that have motivated interest in the phenomenon of background conditions to begin with, would be a negative one.180

Proponents of the deflationary view of background conditions can perhaps get around some of the problems just mentioned by amending it with some additional interpretations. Horty (2012) has recently provided some ideas for how this might be done. On his understanding, background conditions can be explicated by reference to the concept of an exclusionary reason. Horty thinks that some reasons call for responses in the face of other reasons. For example, it may happen that there are reasons to downplay other reasons, or to treat the facts that constitute the reasons as if they were not endowed with any normative status.181 The potential implications that this has for the kinds of examples that we have been looking at are not difficult to discern: The fact that a person is unable to fulfill a promise even in principle might be a reason for him not to consider the fact that he has made the promise to be a reason. What is ignored in that case is not the fact that he has made the promise, but rather the normative status of the fact in question. Similarly, the fact that the colleague is not present to hear a joke could be a reason not to consider the fact that the joke is witty and sarcastic to be a reason for amusement. Once again, what is ignored is not the fact that the joke is witty and sarcastic, but rather the normative significance of the fact.

Horty may be the first philosopher to put explicit emphasis on the relevance of exclusionary reasons in this context. However, careful reflection reveals that the concept of exclusionary reasons is doing some unacknowledged work within

---

180 At the very least, this sounds less reasonable in the case of value than it does in the case of causes. Compare this to the criticism of the pragmatic interpretation mentioned by Cheng & Novick (1991, p. 93). They admit that ordinary people tend to be interested in causal explanations when something uncommon has happened and it is primarily scientists who then go on to identify the prevalent events and describe concepts like viruses and gravity. By itself, this could be taken to speak in favor of the normalcy conditions approach, but Cheng & Novick then point out that ordinary people also understand such concepts when they are used in everyday life and that, in such contexts, their causal status is at least not wholly judged by how commonplace they are in each situation.

181 This could also occur in hurried situations, where a careful consideration of every factor that is relevant to a choice would render a fitting response impossible. Exclusionary reasons, then, may at times aim to simplify deliberation and ensure that this sort of thing does not happen.
the pragmatic interpretation as explicated above. We have seen that on the pragmatic interpretation the background conditions of final value are themselves always parts of value-makers. Equivalently, the background conditions of final value are themselves always parts of the reasons we have to direct attitudes toward that which is good or bad for its own sake. Hence, when it comes to the joke example, background conditions turn out to be parts of our reasons for amusement. When it comes to the promise example, on the other hand, background conditions are instead parts of the reasons to keep the promise that was made. On one understanding, the pragmatic interpretation adds to this the claim that facts about the interests and expectations of speakers, what is commonplace in certain circumstances and conducive to meaningful communication, sometimes constitute reasons of their own. In the kinds of cases that we have been looking at, such pragmatic reasons can even end up being reasons to downplay or disregard the role of other reasons. If this is on the right lines, then it seems to me that we may regard the pragmatic reasons emphasized by the pragmatic interpretation as being exclusionary reasons in a sense very much like what Horyt has in mind.

Both the pragmatic interpretation and the overriding interpretation developed by Horyt involve the concept of an exclusionary reason. What is more, they can both be understood as interpretations of the deflationary view, since they at the very least claim that background conditions must be understood directly in terms of reasons after all. That is not to say that they do not differ in several important respects. The fact that the joke is witty and sarcastic constitutes a reason to be amused by the joke. Both interpretations agree on this. On the pragmatic interpretation, however, the fact that the colleague is present to hear the joke being told is also part of the reason to be amused by the joke. That it is treated as a mere background condition is due to facts about the interests and expectations of speakers—facts that constitute exclusionary reasons. These exclusionary reasons call for the speakers to downplay the normative relevance of the fact that the colleague is present to hear the joke being told. By contrast, we have just seen that the overriding interpretation does not understand the fact that the colleague is present to hear the joke being told as a part of the reason to be amused by the joke. On the overriding interpretation, then, the fact is a background condition only insofar as had the colleague not been present, the fact that he is absent would constitute an exclusionary reason. More specifically, this exclusionary reason would call for speakers to downplay the normative relevance of the fact that the joke is witty and sarcastic.

The overriding interpretation avoids many of the problems that befall the pragmatic interpretation. Because background conditions can now be identified
with pragmatic reasons to ignore other reasons, no real dilemma will arise regarding the concept of intrinsic value. Once it has been couched in terms of the overriding interpretation, the deflationary view does not force us to either reject the first-order position that final value is always intrinsic or reject the first-order view that background conditions can sometimes be extrinsic. Nor will we be forced into situations where we must discount the apparent relevance of a complex of background conditions or introduce that complexity among our value-makers. Finally, amending the pragmatic interpretation with the overriding interpretation in this way also means that we do not have to abandon the substantive intuition that background conditions are not always left out of emphasis because we are not interested in hearing about them. In other words, the overriding interpretation apparently respects the substantive intuition that the presence of the colleague is not part of what makes the joke amusing, just like the ability of a person to fulfill his promise is not part of his reason to fulfill the promise. The biggest problem with the overriding interpretation, and the reason that I see it as a deflationary view, is that it understands background conditions directly in terms of reasons. It claims that a factor constitutes a background condition if, and only if, its absence would constitute an exclusionary reason. I submit that an account of background conditions would be preferable if it did not make this commitment.

As a final critical note, it should be pointed out that the reading above takes some liberties when it comes to the views of Horty. In his original discussions of exclusionary reasons, he explicitly denies that exclusionary reasons only call for us to downplay the normative importance of other reasons. There is some unclarity here, but Horty seems to think that when exclusionary reasons call for us to downplay the normative importance of other reasons, those other reasons also lose their normative importance (Horty, 2012, p. 128). Whether this is somehow the result of our downplaying the reasons or the result of the facts being present which constitutes the exclusionary reasons is unclear. Regardless, this is evidently a far more robust account of background conditions than a proponent of the deflationary view would be comfortable with. What speaks against the robust reading of Horty in this context is that it leaves the fundamental questions with which we are currently engaged largely unanswered. To make this clear we can restate the questions here in a more specific form: If it is true that a reason can lose its normative importance as a result of there being exclusionary reasons or other facts present, then what is the nature of this effect? What do we commit ourselves to when we judge that one reason has snuffed out another in this robust sense? We will momentarily return to some of the issues that have just been explored.
Beyond the Deflationary View

I can find consolation in the fact that there are a few philosophers who have taken an interest in background conditions and who do not accept a deflationary view. Bader (2016) is among them as he tries to show that background conditions should be taken more seriously than the deflationary view tends to do, presumably irrespective of whether it takes the form of a pragmatic interpretation or overriding interpretation. His objections to the deflationary view are also refreshing in that they are not meant to rely solely on appeals to substantive examples but instead gain most of their weight from observations about *absences of disablers*. To reiterate, disablers ensure that factors that would otherwise have been effective value-makers become normatively inert. Disablers are in this sense meant to do the opposite job from enablers but are still thought to involve the same general kind of phenomenon. Bader focuses on background conditions for reasons and argues that complete normative explanations for the presence of reasons will not only refer to the presence of some factors but will also refer to the absence of other factors. Such absences cannot possibly be part of the source of reasons, he claims, but may well be part of the background conditions of reasons (ibid, p. 32). Bader thinks that the significance of absences of disablers and the possibility that they play the role of background conditions is straightforwardly revealed by counterfactual considerations:

If the condition fails to be satisfied, i.e. if the thing in question is not absent but present, then there is something that can play a role and that can prevent the source from constituting a reason, namely a disabler. The absence is accordingly relevant, insofar as, were it not for this absence, the source of the reason would have been disabled, such that it would not have given rise to the relevant reason. Accordingly, what makes it the case that the consideration constitutes a reason is not that there is something that is present and that plays a positive role in making it a reason, but that certain things are absent that would prevent it from constituting a reason were they to be present. That is to say, absences are only relevant in that the reason is merely negatively dependent on them (ibid, p. 32).

I find that there are several distracting issues that make this argument hard to follow. One of the issues concerns the precise role that Bader wants to ascribe to absences of disablers. If something has a contribution to normative explanations,
then it has a contribution that is either positive or negative in character. Bader does not regard absences of disablers as being disablers themselves, since they are not thought to have a negative contribution to complete normative explanations. This suggests that he instead understands absences of disablers as enablers, but he later denies this as well. Bader (2016, pp. 33–35) argues that the commonplace view that absences of disablers are equivalent to presences of enablers is incorrect. He suggests that whereas enablers and disablers are to be understood in terms of the effect that their respective presences have on reasons, absences of disablers should be understood in an entirely separate way. Bader says that absences of disablers are a counterfactual phenomenon, explained in terms of the effect that factors would have had if they were instantiated. This does indeed suggest that absences of disablers cannot play the role of reasons, but, contrary to what Bader is hoping, it also suggests that absences of disablers cannot play the role of background conditions. That something stands in some form of counterfactual relation to reasons for attitudes is not sufficient for it to be regarded as a background condition.

This criticism of Bader gains support from the observation that other factors, such as contributory background conditions mentioned at the start of this chapter, can also be counter-factually relevant to the presence of reasons. Another issue that I have with Bader’s argument concerns his views about facts about absences. Even if Bader is right to suggest that absences as such cannot constitute reasons, it is much less clear that facts about absences cannot constitute reasons. Bader explicitly denies the validity of this line of reasoning in a footnote (ibid, n. 10), but he does so without much argument and by appeals to intuitions that I just do not share. That there is an absence in ability to fulfill a promise may be a fact. Philosophers who accept the deflationary view could regard said fact as a reason itself or as part of a reason. For example, following the understanding developed by Horty, the fact that an ability to fulfill a promise is absent might be an exclusionary reason to ignore that the promise has been made. The same general point can be made in terms of value-makers as well. That a colleague is absent when a joke is being told at her expense could be a fact. Philosophers that accept the deflationary view might regard the fact as a value-maker in respect of the joke and its amusing nature. For example, following the understanding of Strandberg and Fogal, the fact that an ability to fulfill a promise is absent might be part of the reason, or normative cluster that recommends us to be amused.

182 However, Bader acknowledges that “the presence of an enabler implies the absence of disablers, and the presence of a disabler implies the absence of enablers” (ibid, p. 34).
Proponents of the deflationary view can respond to the objections offered by Bader rather straightforwardly, by insisting that the counterfactual considerations he mentions are misleading when taken at face value. Whenever counterfactual considerations seem to indicate that absences as such can help explain the presence of reasons, relevance should instead be ascribed to facts about those absences. This explanation would not be *ad hoc*, since even skeptics of the deflationary view such as myself might agree that there is something intuitively strange about ascribing the role of background conditions to absences as such. Background conditions are meant to be constituted by facts, properties, relations, and the like—not the failure of these things to be instantiated. That this is an aspect of background conditions that is not made clear by counterfactual considerations alone is not very surprising, nor can this fact itself be regarded as particularly compelling evidence against the deflationary view. Indeed, evidence in either direction of the debate seems thin on the ground. My view is that the main reasons speaking against the deflationary view of background conditions are less technical.

It is a reasonable rule of thumb that philosophers should try to be as accommodating as possible when it comes to different first-order views. This means that we should prefer an account of background conditions that retains the relative parsimony of the deflationary view, while doing a better job at capturing the substantive intuitions surrounding the phenomenon. Only when we have failed to formulate such a conciliatory account, or else established that the substantive intuitions are misleading, should we consider putting the first-order views on the sacrificial altar. It is also relevant here that no matter how detailed or sophisticated the deflationary view is made, there will probably remain enough conceptual wiggle room so that we can deny that any given background condition is part of any reason.\(^{183}\) This does not necessarily show that the deflationary view is unsound but only that it does not get us as far as we might like. It may even be argued that the deflationary view amounts to little more than a kind of skepticism on which background conditions do not have the effect that we may have supposed. This could be true, but the question

\(^{183}\) The empirical studies of Cheng & Novick (1991, pp. 101–117) indicate that even though the pragmatic interpretation may accurately predict what will be left out of emphasis in causal explanations, it does not predict what factors speakers and hearers judge to be causes when pressed on the issue. This has yet to be shown to be true in the normative case as well. However, if it is true in the normative case, then this would give weight to the view just expressed, that the deflationary view only explains the pattern by which we leave things out of emphasis in normative explanations, but not what factors we judge to be background conditions or indeed what we commit ourselves to when we make such judgments.
remains what it might mean to think or talk of something as if background conditions did have this effect.  

Besides offering objections to the deflationary view, Bader also presents many positive remarks on the nature of background conditions. Although these remarks are often plausible on their own, it can be difficult to discern how they are meant to go together, and to ascertain the extent to which they might come into conflict with the account soon to be defended in this chapter. Let me first make some comparisons regarding terminology, for even in this regard there seem to be some discrepancies that make interpretation of a bit difficult. In the sources where Bader is mainly concerned with the background conditions of reasons, he begins his discussions by emphasizing the distinction between “that in virtue of which something is a reason (the source or ground of the reason)” and “that on condition of which something is a reason (the conditions of the reason)” (ibid, p. 28). Bader later clarifies his emphasis by stating: “Whereas the source or ground of a reason is to be identified with the consideration that constitutes the reason, conditions are not themselves part of the reason” (ibid, p. 31). This way of phrasing things does not fit well with the manner of speaking that has been employed here, and Bader’s emphasis on the distinction between the facts that constitute reasons and their conditions is something that I will have to say more about later. The view that I will be defending states that conditions of final value are the factors in virtue of which, or as a result of which, facts have the normative property of being a reason. In regard to the question how the background conditions for reasons should be understood, it will turn out that we have at least two different options. In any event, Bader also offers the following positive characterization of background conditions:

---

184 An analogy could perhaps be made here with the error theory of value as it occurs in the writings of J. L. Mackie (1977) and Richard Joyce (2001). Both philosophers think that value is not a genuine property or relation in the world. They also provide pragmatic stories about why we talk as if this were the case. However, they do not stop there. They also trouble themselves with saying something about what kind of thing value would be if it existed. They carefully consider the conceptual issue of what is involved in judging something to be good or bad and, in fact, base much of their skepticism on the resulting observations. Among them is the observation that value is represented in our normative talk and practices as being too “queer” to be taken seriously (Mackie, 1977, pp. 38–42). Proponents of the deflationary view and its pragmatic interpretation have yet to do anything similar in the case of background conditions. For philosophers interested in conceptual analysis this should itself be a source of dissatisfaction, but beyond this it at least weakens whatever metaphysical claim they have against existence of background conditions.
These cases show that the properties that are selected at the non-normative level do not by themselves suffice for grounding normative properties. Since they are only conditional grounds, they only have normative significance conditionally. Accordingly, one has to appeal to the normative level to identify and explain the conditions under which these features do in fact give rise to normative properties, namely when the enabling/disabling conditions are satisfied. This means that these non-normative features are only to be selected under certain conditions, where the selection is determined by the normative grounding principles. (2017, p. 132).185

Bader says that value-makers are conditional and that their conditions are what we speak of as being background conditions. The same is meant to apply to reasons, but what is conditional in that context is the fact that certain other facts are reasons. Now, Bader also insists at one point that enablers of enablers are impossible, “given that enabling is not a grounding relation that can be subject to conditions” (2016, p. 38). When taken together, all this could be taken to mean that background conditions do not involve grounding relations at all and so cannot be subject to conditions of their own. Perhaps a fairer interpretation takes Bader to mean that the specific type of grounding relations that are involved in background conditions cannot be the type that is subject to conditions. This would be just another fundamental respect in which value-makers differ from background conditions. Clearly, this does not fit well with my previous comments about the possibility of there being enablers of enablers, but it might still be compatible with the view that I will ultimately be defending.

185 As with many of Bader’s comments and observations about background conditions, there is a lot going on here in a relatively small space. In the first place, there seems to be an idea being expressed here about whether dependence can include only nonnormative factors. Bader seems to suggest that this is not possible, since background conditions, which are included in dependence according to our present understanding, cannot be understood in wholly nonnormative terms. I find this line of reasoning somewhat perplexing as well and I wish to briefly mention why. Dependence can include nonnormative factors, since even background conditions can be constituted by nonnormative factors. Of course, this does not mean that the role of being a background condition can be understood in wholly nonnormative terms, but this observation seems rather trivial in the current context. Value-makers can also be constituted by nonnormative factors and yet the fact that the role of being a value-maker cannot be understood in wholly nonnormative is obvious. With regard to these issues, then, it seems to me that what Bader is saying is either trivial or wrong. Be that as it may, all this is merely a distraction from the main takeaway here, which is that Bader understands background conditions in terms of the rather controversial concept of conditional grounding.
Admittedly, other terminological and theoretical discrepancies make it difficult to establish with any degree of certainty how much Bader would sympathize with that view, so all I can do for now is hope that we are largely in agreement. However, given the tight connection that he posits between the understanding of his background conditions and counterfactual reasoning, I must admit to skepticism here. It is important to remember that the account soon to be presented has a strong hyperintensional flavor and furthermore avoids appeal to the controversial concept of conditional grounding. More specifically, it will turn out that if we have a view on which a normative property, like final value, can be understood in terms of the presence of reasons, then the background conditions of the former can be analyzed without appeal to the notion of conditional grounding. Regarding the background conditions of reasons themselves we either have the option of denying the distinction between background conditions and makers, or we accept the notion of conditional grounding. I will say more about this in a moment.

**Toward a Hyperintensional Analysis**

The purpose in going into the theoretical details of the deflationary view was to provide us with some background and to hint at the difficulties that are involved in the application of the view within the normative domain. If there is any lesson to draw from the last few sections it is mainly that philosophers must not be too quick about endorsing the deflationary view of background conditions, unless they are also willing to put in the work needed to make it feasible. For example, if it is true that background conditions are just pragmatic tools, not fundamentally different from value-makers or reasons, then we need some account of how they are used: Is there a discernible pattern to our thinking and talking about background conditions? What exactly determines whether a factor is left out of emphasis in a normative explanation? Whenever the pragmatic interpretation is offered as an explanation of the concept of background conditions, these tough questions will need to be answered. We have seen that trying to answer the questions might present problems that were not present in quite the same way when the deflationary view is applied to the case of cause and effect. What is more, whether it can be made to work or not, it seems reasonable to assume that the deflationary view can only be part of a wider explanatory program dealing with background conditions. Here we shall therefore consider what else can possibly be said about the concept in question.
The dependence of final value is a broad relation that is constituted by several narrower ones, among them being the relation of evaluative resultance. The dependence of final value and evaluative resultance were described in the introduction as having common characteristics, the most central of which were their reliance upon a thick notion of explanation and priority. These characteristics were brought out by considering how normative explanations, of both the wide and narrow kind, seem to impose a metaphysical structure onto the world, whereby one thing is deemed to be more basic than another. For example, when the final value of an object is judged to depend upon, or evaluatively result from, the presence of a given property, the latter is seen to be responsible for the former in a sense that cannot be couched in mere extensional or counterfactual terms. Now, it seems reasonable that we should apply this approach generally and suggest that all the factors that play a significant role in broad explanations of final value are of this sort, including those that are picked out by the concept of background conditions. After all, judging that a fact enables the final value of an object also means imposing a metaphysical structure onto the world. Perhaps this needs to be accommodated as well.

Before I set out the account I wish to defend we should also reiterate some of the observations made in the introductory chapter regarding the nature of reasons. Whenever there are reasons for responses, facts must obtain that constitute the reasons in question. Facts should in turn be understood as obtaining states of affairs, which are entities with a propositional character that can be picked out by that-clauses. That a person has promised a friend to water her plants is a state of affairs and once it obtains it becomes a fact. Only then can it have the normative property of being a reason. We should also recall the distinction between reasons and the normative property of being a reason. Whenever certain facts constitute reasons, then, there must be factors that make them reasons. Just as objects cannot have value without there being something that makes them good or bad for their own sakes, facts are always reasons in virtue of something else. When these observations are looked at through the lens of the fitting-attitudes analysis, we start to see that what makes an object have final value is always a fact. What makes an object good or bad for its own sake is the fact that constitutes the reason we have to direct a positive attitude toward the object. Language permits apparent ascriptions of the value-maker role to other things, including the properties that a value-bearer has, but the role is really being played by the fact that the value-bearer has those properties.

The account I wish to defend says that the background conditions for final value are not the makers of value, but rather the things in virtue of which, or as a result of which, facts are endowed with the normative property of being a
The implications that this account has regarding the first example mentioned in the introduction are easy to illustrate: A joke is being told at the expense of a colleague, whose presence is required for the joke to be amusing. Yet, the presence of the colleague does not seem to be among the makers of the relevant value. In this case, we could instead judge the presence of the colleague as being part of what makes the fact that the joke is witty and sarcastic a reason to be amused by the joke in the first place. The fact that the joke is witty and sarcastic is a reason in virtue of the presence of the work colleague. The substantive intuitions invited by the example, on which background conditions are completely distinct from the makers of value, is thereby captured. The background conditions for the final value are instead identified as being part of what provides the fact that the joke is witty and sarcastic with its normative status. The account also appears to me theoretically parsimonious, as it does not commit us to a new type of relation. At the very least, it does not saddle us with more mystery than the normative domain is already burdened with. The notion of conditional grounding that Bader has formulated is also not needed.

Let me now also address the background conditions for reasons, for here things become slightly more complicated. The fact that one person can fulfill his promise to his friend that he will water her plants seems relevant to the question whether he has a reason to water her plants. Yet, the fact in question is not itself necessarily a part of any reason that might be relevant to the case at hand. The reason he must fulfill his promise consists solely of the fact that he made the promise to begin with. Now, there are two ways in which my account could be spelled out here, depending on what stance we take regarding an apparent asymmetry in the observations made by Dancy. He takes the background conditions of value to be completely distinct from the makers of value. The background conditions of value contribute to other factors playing the role of value-makers, without being parts of the value-makers themselves. By contrast, Dancy takes the background conditions of reasons to be distinct from the facts

---

186 The distinction between reasons and the normative property of being a reason is also relevant to the understanding of traditional theories about the nature of normativity, including subjectivism. On the standard view, subjectivism and its rival objectivism disagree about what facts could be said to constitute our reasons. The story goes that while subjectivism claims that reasons are only constituted by facts about our attitudes, objectivism denies that this is so. If what has been said so far is right, then there is an alternative understanding on offer. On this understanding subjectivism and objectivism disagree about in virtue of what certain facts are reasons. The fact that a storm is coming might be a reason for us to go outside. The fact that a friend is sad might be a reason for us to comfort her. The fact that smoking is unhealthy might be a reason for us to quit. What the theories disagree about is rather the issue of how to best explain why this is so. This interpretation is explored in the next chapter.
that constitute the reasons. He is therefore not as careful to distinguish the
background conditions for reasons from reason-makers as he is to distinguish the
background conditions for value from value-makers.

I suspect that this apparent asymmetry is not simply an oversight on part of
Dancy. Rather, he seems to take the background conditions for value to have a
more nuanced structure than the background conditions of reasons. The fact
that one person can fulfill his promise is a reason-maker, meaning that it
constitutes not a part of the reason he has to fulfill his promise, but rather the
factors that afford the fact that he has made it its normative status. On this
account, which fits very well with the apparent asymmetry in Dancy, the
background conditions for reasons are understood not to be distinct from
reason-makers. The alternative would be to say that the fact that the person is
able to fulfill his promise affords some other factor the role of being a reason-
maker. There may be doubts about whether there are any such factors, since
substantive examples seem rather thin on the ground. If the capability of the
person to fulfil his promise affords other factors their role of being reasons-
makers, then what are those factors? What is here providing the fact that he
made the promise with its normative status? It is also important to note that if
we take Dancy at his word and assume that there is an asymmetry between the
value case and the reason case, then this seems to play right into the hands of
proponents of the pragmatic interpretation. Strandberg, who was identified an
enthusiastic defender of this approach, at one point writes:

On this view, it is possible to account for the kinds of case Dancy
appeals to in his argument. We might maintain that what
constitutes a person’s reason to perform an action is that it is such
that she has promised to do it, not a complex of non-moral
properties which includes a non-moral property which Dancy
classifies as an enabler, such as the action being such that she freely
promised to do it. Hence, a non-moral property which Dancy
classifies as an enabler (e.g. an action not being such that the
promise to perform it was given under duress) is not part of a
reason to perform an action. However, this view leaves open the
possibility that although such a non-moral property is not part of
a reason, it is still relevant to the nature and the strength of the
reasons a person has to perform an action (2008, p. 147).

Clearly, even the most enthusiastic proponents of the deflationary view
should be able to admit that reasons have their makers, and so they could easily
admit that reasons have background conditions \textit{in that sense}. If we want to understand the background conditions for reasons in a more robust way, then we will have to accept that there are makers of reasons-makers, which seems to bring us very close to arguing what Bader argues when he talks of conditional grounding. I think that all of this points to an interesting general observation regarding background conditions and the structure of normative explanations. If normative properties are generally understood in terms of the presence of reasons, then we can explain what it means for them to be sensitive to background conditions without much strain, by invoking the factors that afford facts with the normative property of being a reason. However, reasons themselves, insofar as they are treated like the normative bedrock on which everything else stands, are perhaps a bit different. Here it may be more reasonable to say that while background conditions are not to be conflated with the facts that constitute reasons, there is really no need to distinguish background conditions from reasons-makers.

At the risk of disappointing my readers I wish to remain neutral on some of the issues just raised. Irrespective of what stance we take regarding the asymmetry, my account of the background conditions for final value manages to be sufficiently accommodating of the substantive intuitions we have so far focused on and it does not rely on the notion of conditional grounding. Before we move on I would like to consider some objections to that account. One such objection says that all manner of things can be part of the explanation why a fact is afforded its normative status, but not all of them are background conditions for the corresponding final value. For example, perhaps it makes sense to say that the fact that a joke is witty and sarcastic is itself part of the explanation for why that very fact ends up being a reason to be amused. Even if this is right, we may want to deny that the fact that the joke is witty and sarcastic, along with the presence of the colleague, is an enabler. One response would be to go more fine-grained here and suggest that background conditions are those specific parts of the dependence base for the normative property of being a reason that are themselves not part of the reason. This is enough to exclude the fact that a joke is witty and sarcastic from being considered an enabler in this example.

Specifying that background conditions are the factors that afford facts the normative property of being a reason without being parts of the facts themselves also helps explain why background conditions are not essential parts of normative explanations. For a given reason it might be that the fact that constitutes it is afforded its normative status only by the presence of a factor which is itself part of the reason. In that case, the reason, and the value to which it might correspond, does not depend upon background conditions. The
requirement under consideration therefore saves us from the risk of an infinite regress of background conditions. The worry is that once we admit that if normativity always has a source and it is possible for there to be such things as enablers of enablers, we will be forced to see the chains that constitute complete normative explanations as potentially being infinitely long. The requirement we have just been considering helps us avoid the problem since it allows us to say, at any point in an explanatory chain, that a factor is afforded its normative status by its own inherent nature. It follows that it is not a background condition.

The account assumes that there is a distinction between the questions what makes an object good or bad for its own sake and what makes a fact a reason for attitudes. This assumption is used to accommodate the intuition that the makers of final value and the makers of reasons for attitudes are different things. Asking what makes an object have final value is the same thing as asking what reasons there are to direct a certain kind of attitude toward the object. To my mind, this is an existential question which calls for an answer in terms of a certain fact. The further question why this fact is a reason to begin with is a different question altogether that represents a move away from talking about value-makers. However, it seems as if there is an opportunity to push back here and insist that the questions cannot be meaningfully distinguished after all. The joke is amusing in virtue of the fact that it is witty and sarcastic. This is to say that the fact that the joke is witty and sarcastic is a reason to be amused. The fact in question is a reason to be amused in virtue of the fact that the colleague is present. The transitivity of the relevant sense of explanation thereby guarantees that we are committed to saying that the joke is amusing in virtue of the colleague being present as well.

The objection just described would be very serious if it turned out to be sound, because it would entail that the account I have developed would fail to be as accommodating regarding substantive intuition as it is meant to be. Not surprisingly, I do not think that it is sound. In fact, I suspect that the objection trivializes the deflationary view. It cannot be the case that proponents of the deflationary view only suggest that background conditions for final value must be part of its explanation from a very general perspective, and that from this general perspective, background conditions are not so different from value-makers. For if this is all that proponents of the deflationary view mean to claim, then it is difficult to see how it would be possible to disagree. I have so far been very careful to distinguish between evaluative resultance and dependence and characterized it as part of the very meaning of the concept of a background condition that its presence contributes in some way to the instantiation of value. For example, a factor is assumed to play the role of an enabler only if it takes
part of normative explanations in the very wide sense that encompasses dependence. It would be strange to suggest that by assuming this much I have already admitted that proponents of the deflationary view are right. However, it is one thing to say that this objection must get something wrong and it is quite another to explain exactly what the objection gets wrong.

Invoking the distinction between evaluative resultance and the dependence of final value will not do as a response, because such an approach risks trivializing the account that I want to defend. The deflationary view, at least according to its pragmatic interpretation, claims that the distinction between value-makers and background conditions is merely a pragmatic distinction in language. If my response to the objection just mentioned ends up relying upon a distinction between evaluative resultance and dependence of final value, which might itself appear to be rather stipulative, then one can begin to wonder how much is really won. Now, one of my main disagreements with the deflationary view concerns the question what should be regarded as value-makers to begin with. In its pragmatic interpretation, the deflationary view identifies the presence of the colleague as part of what makes the joke amusing. I have presented objections toward this claim by pointing out that the fact in question is not a constituent of the fact that constitutes the reason to be amused by the joke. Insofar as the presence of the colleague is relevant, it must instead play the role of a background condition that makes the fact that the joke is witty and sarcastic a reason to be amused by the joke.

The response just given is obviously compatible with the suggestion that the fact that the colleague is present is part of the big picture which we would have to paint in order to give a complete explanation for why the joke ends up being amusing. Indeed, the response just insists upon some of the stipulations made in the introductory chapter by stating that what should be regarded as value-makers are only the things found in the resultance base for final value. However, I wish to emphasize that this is not a mere terminological matter. If an object is good or bad for its own sake, then its value-makers are the factors that constitute the reasons we have to direct attitudes toward the object. Other things may be relevant to complete explanations, to be sure, but they are not makers in the sense that we have in mind when we ordinarily think and talk about final value. In such contexts, the factors that we identify as value-makers are specifically those that are important to questions about how we should respond to the objects and choices that life presents us with. In this way, the distinction between value-makers and background conditions tracks something important about the way final value relates to other things.
Not only do I find the fitting-attitudes analysis of value to be intuitive, but I consider it to be the jewel of value theory. I have already relied upon the fitting-attitudes analysis of value quite a lot and described it in the introductory chapter as an essential part of the framework which was meant to be applied within the book. Nevertheless, I would like to consider the worry that the general applicability of the account I have here defended would be lost if value cannot be understood in terms of reasons for attitudes. Even if the fitting-attitudes analysis of value is correct, it might be argued that the account I have been defending fails in respect of theoretical economy by its reliance upon the analysis. At the very least, an account of background conditions that did not have to rely upon the fitting-attitudes analysis of value to make sense would be preferable. In response to this argument, I wish to point out that the fitting-attitudes analysis of value enables me to account for the background conditions for value and possibly the background conditions for reasons in a unified way. Indeed, depending on how we regard the apparent asymmetry in Dancy, the background conditions for final value might turn out to be the very same factors as the background conditions for reasons for attitudes.

The unified picture of the background conditions of value and the background conditions for reasons would be lost in the absence of the fitting-attitudes analysis of value. I admit as much and make no apologies for it. However, it is important to note that the hypothetical failure of the fitting-attitudes analysis of value would not affect how the account could be applied to the background conditions for reasons. The only change in that respect is that once the application to the background conditions for reasons has been carried through, we would still need a separate story about the background conditions of value. Of course, it is also true that if we are forced to accept a deflationary view of the background conditions of value, the account of the background conditions of reasons suggested by Dancy’s comments could help make the deflationary view more plausible. Substantive intuitions suggest that the presence of the colleague is not part of what makes the joke amusing, for what plays this role is rather the fact that the joke is witty and sarcastic. To explain away this intuition, we could now also suggest that although the presence of the colleague is part of what makes the joke amusing after all, it nonetheless serves as a background condition for the reasons we have to be amused by it. In this way, the behavior of reasons can affect the intuitions we have about the value case.

Finally, I wish to say something now about the concept of a derivative value, which is often used by philosophers but seldom properly explicated. Suppose that a parent has reasons to admire a painting for its own sake because it was made by her children. More specifically, the fact that the painting was created by
her children constitutes a reason for the parent to admire it for its own sake. It follows that the painting is admirable in a final sense. The admirability of the painting might be judged to be derivative because it is parasitic on the presence of other reasons, namely the reasons that the parent has to promote the wellbeing and creative development of her children. This seems readily explainable within the sort of account that has just been developed. If a fact has the normative property of being a reason and this is in virtue of, or a result of, another fact having the same property, then the latter fact ends up corresponding to a derivate value. In this specific case, the fact that the painting is made by the children of a parent is endowed with its normative status by the reason that the parent has to promote their wellbeing and creative development.

Of course, it is probably not uncommon that when we judge objects to have derivative value, we do not really judge them to be good or bad at all. Instead, we judge the objects to be good or bad only so that we may hint at the presence of some other object which is a bearer of value. In other words, we sometimes use the concept of derivative value to describe certain objects only to indicate that something else is good or bad. When a hedonist utilitarian judges that certain possible worlds have greater final value than others, she presumably has in mind a derivative value of this merely indicative kind. For what she means in such a case is that certain possible worlds contain more things that are good or bad for their own sakes than others. However, there appears to be some tension between this use of the concept of derivative value and the fitting-attitudes analysis. For if the hedonist utilitarian judges that we have reasons to prefer certain worlds for their own sakes in virtue of, or as a result of, their containing more pleasure than others, then, given the fitting-attitudes analysis of value, she seems committed to judging those worlds to have final value. The story laid out in the previous paragraph suggests one way of avoiding the problem.

The idea is that possible worlds have derivative value in a slightly different sense than the merely indicative one that was just described. The fact that certain possible worlds contain more pleasure than others constitutes a reason for us to prefer them for their own sakes, but its normative property of being a reason depends upon the presence of yet another reason. In this way, the final value that is ascribed to certain possible worlds is less fundamental, or less robust, than the final value which accrues to pleasure itself. For it is ultimately in virtue of, or as a result of, our reasons for holding positive attitudes toward pleasure for its own sake that the fact that certain worlds contain more pleasure than others is afforded its normative status. Whether a hedonist utilitarian would be entirely happy with this line of response is uncertain, but it seems to me to rely precisely on that concept of derivative value which is invited by the
account I have just been defending. This concept, which gains support from assumptions concerning the distinction between reasons and the normative property of being a reason, shares certain features with the commonplace understanding of instrumental value. Let me say something more about this.

Derivative value is like instrumental value in that it is a rather cheap type of value. All manner of things might have derivative value, but their being good or bad for their own sakes is to be partly explained by reference to there being reasons to hold attitudes toward something else. And at the end of the explanatory chain there should be a nonderivative final value. If this is right, then first-order theories about value, like hedonist utilitarianism, should perhaps be exclusively recognized by what they identify as having this kind of final value, rather than the things they claim to be derivatively good and bad for their own sakes. This would also entail that the concept of background conditions, as the account currently under consideration understands it, might not be such a peripheral concept as we may have been led to believe. In fact, substantive examples of background conditions can be found all over the literature, in many of the instances where philosophers have said that one final value is derivative or somehow parasitic upon some other normative consideration. Although background conditions play a less direct role than value-makers do in the context of normative explanations, then, that is not to say that they are not also very central to our evaluative talk and practices.

I wish now to recapitulate briefly and put what has just been said into perspective. Final value is not only affected by its makers for it is also affected, in a less direct manner, by the presence of background conditions. This section has been dedicated to developing an account of this concept, which aims to explain what sort of effect background conditions are general meant to have on final value. On the account in question, the background conditions for final value are the factors that explain why facts are endowed with the normative property of being a reason. The background conditions of final value are to be explained, then, not as being parts of value-makers, but as being parts of reasons-makers. When it comes to the background conditions of reasons, these are either to be seen as reasons-makers as well, or as the factors that make other factors into reasons-makers. Before this section ends, I wish also to point out that the concept of background conditions is plausibly wider than we sometimes hope that it is. What this means is that the concept might pick out a large category of heterogeneous phenomena, only some of which are captured by my account. That account can help us understand many substantive examples of background conditions, but whether it can help us explain all of them is something that I wish to remain silent about.
Concluding remarks

To understand how final value depends upon other things, we need to look beyond evaluative resulance and investigate the other phenomena which occur within the dependence of final value. This chapter endeavored to do so by attempting to clarify the concept of background conditions. Factors playing the role of background conditions help make narrow normative explanations correct, without themselves being part of them. More precisely put, background conditions explain how other factors can play the role of value-makers or reasons, without themselves thereby being either value-makers or reasons. The question explored throughout the preceding discussion is what else background conditions could possibly be: How should we understand the role that is played by a factor when it is judged to be a background condition for final value? The chapter began by making some very general observations about the nature of background conditions, with the aim of distinguishing them from other relations that have a similar logical or counterfactual relationship to the presence of final value. This included contributory background conditions which help why objects are in possession of value-makers in the first place. We also set aside epistemic background conditions which help explain how we can become aware of the value-makers that objects have.

A large part of the chapter was then dedicated to discussing the deflationary view of background conditions, with special attention paid to the pragmatic interpretation. This interpretation, which derives its inspiration from a parallel view within the philosophy of science, insists that background conditions are either value-makers or reasons themselves. The fact that background conditions are often treated as something other than value-makers or reasons is, as the name of the interpretation implies, mainly to do with the pragmatics of our evaluative talk and practices. I argued that the pragmatic interpretation is more difficult to apply to the case of final value than appearances might suggest, but that it can be made more plausible by amending it with some recent insights concerning exclusionary reasons. According to the overriding interpretation a factor is a background conditions in so far as had the factor not obtained, then this would constitute a reason to downplay or ignore the presence of other reasons. Amending the pragmatic interpretation with the overriding interpretation appears to give us a more flexible deflationary view, but I argued that we nonetheless ought to put the deflationary view aside here.

After discussing some of the problems associated with the deflationary view, I moved on to consider the hyperintensional analysis of background conditions that I wanted to defend. On this analysis, the background conditions of final
value ought to be described in terms of the factors in virtue of which, or as a result of which, facts become reasons for attitudes. Whenever facts are reasons for attitudes, then there is some factor that explains why this is so. This factor would constitute a background condition in respect of the resulting final value. Having detailed the implications of my analysis I then responded to some objections, some of which allowed me to make some additional clarifications regarding the proper understanding of value-makers and their role in normative explanations. Finally, the fruitfulness of my analysis of background conditions was illustrated partly using substantive examples and partly by showing how it can be applied to the concept of derivative value. The idea, roughly, was that derivative value corresponds to reasons for attitudes that result from the presence of another reason. In other words, when the instantiation of the normative property of being a reason in one fact helps explain why another fact is also afforded that same property, then what results is a derivative value.

The account I have been defending was meant to be supported by some of the assumptions mentioned and defended in the introductory chapter. What is more, the account in question had two separate theoretical aims: The first aim was to be able to capture the sort of substantive intuitions that were invited by the standard examples of background conditions. These substantive intuitions suggest that background conditions cannot themselves be ordinary value-makers or reasons; the second aim was to retain the theoretical economy and parsimony of the deflationary view and avoid having to explain background conditions in terms of an entirely new relation. I argued that the hyperintensional analysis seems to fulfill both aims. Nonetheless, I also conceded that there may in the end be a need to acknowledge that there is a plurality of background conditions. The phenomenon of a factor contributing to the applicability of a certain narrow normative explanation without taking part of that very explanation may involve a more heterogeneous one than previously proposed. If this is right, then what I have done is clarified one way in which talk of background conditions can be made intelligible.
Chapter 5
On Constitutive Grounds

When exploring the dependence of final value upon other things, philosophers have recently adopted a wider perspective that looks beyond the phenomenon of evaluative resultance. The result is that things like background conditions and constitutive grounds have received increased attention in the literature. The previous chapter introduced and discussed the concept of background conditions in closer detail. Its purpose was, in part, to defend an account the aim of which is to describe background conditions of final value in terms of the factors that make facts reasons. The following can be viewed as a companion piece to that chapter, in so far as it aims to do roughly the same thing when it comes to constitutive grounds.187 However, the chapter goes about doing this in a somewhat roundabout way, by paying close attention to a disagreement between two rival theories about the nature of final value: The first theory is that of subjectivism, which asserts that final value is necessarily determined by the presence of attitudes. According to this theory, if an object is not targeted by an attitude, then the object cannot have final value.188 The second theory is that of objectivism, which denies that final value is necessarily determined by the

---

187 Much gratitude is once again owed to Jakob Green Werkmäster for discussing these issues with me and helping me develop the main ideas and objections found within this chapter. Some of the sections to follow occur in an earlier form in a paper we have collaborated on and which was recently published. See García & Green Werkmäster (2018).

188 Whether the attitude must be present in the actual world or in a possible one in which the object exists and has value is a tricky question to which we shall return in a moment.
presence of attitudes in this way. According to this theory, an object can have final value, even if it is not targeted by an attitude. Exactly how the disagreement between the two theories just mentioned should be understood is a complicated issue, to which the concept of constitutive grounds is sometimes thought to be of help. Consider the following:

**Value Constitution:** Final value accrues to objects only to the extent that the objects are either favored or disfavored for their own sakes. Attitudes themselves are not bearers of final value, unless they too are among the things that are targeted by attitudes in this way.

**Maker Constitution:** Final value accrues to objects in virtue of factors only to the extent that the objects are targeted by attitudes in which the factors are represented as value-makers. Attitudes themselves are not value-makers, unless they too are represented in this way.190

When subjectivism insists that attitudes serve as the constitutive grounds for final value, it holds that both claims are true. When objectivism denies that attitudes serve as the constitutive grounds for final value, it holds that both claims are false. We can also imagine an intermediary position, which holds that only one of the claims is true and the other is false.191 This way of understanding

---

189 Subjectivism and objectivism arguably differ in terms of their inspirations and historical roots as well. The former theory has sprung from a tradition which denies that final value is a genuine property or relation in the world. The latter theory has sprung from the opposite tradition of affirming that final value is a genuine property or relation in the world. Although much of the following discussion will be quiet on this ontological issue, we shall see that there are instances where it is relevant to the understanding of the framework of constitutive grounds. On a related note, it should perhaps also be acknowledged that some philosophers have understood subjectivism to be a theory on which preference satisfaction is good or what we have most reasons to promote (e.g. Moore, 1922; Korsgaard, 1983; Kagan, 1998; Parfit, 2011). Other philosophers take subjectivism to make claims about the substantive importance of wellbeing and what is of personal benefit to a person (Sobel, 2017). It is also part of the commonplace to understand subjectivism as a view claiming that our value statements either express or give nonnormative reports of the speaker’s attitudes (e.g. Hare, 1981; Rachels & Rachels, 1986; Bar-on & Chrisman, 2009; Jackson, 1998; Miller, 2003; Shafer-Landau, 2003). We shall put aside such understandings of subjectivism here.

190 Because we are assuming that there can be no value without value-makers, proponents of subjectivism should claim that as a matter of conceptual necessity, value constitution cannot take place without maker-constitution also having taken place. See Fritzson (2014, pp. 84–92).

191 Perhaps subjectivism and objectivism do not exhaust the domain of possible theories here. Some philosophers would tend toward constructivism, which could be seen as a third and independent alternative. Constructivism will be left aside here for two distinct reasons. The
things represents a framework for thinking about the disagreement between subjectivism and objectivism: the framework of constitutive grounds.\textsuperscript{192} The primary benefit of the framework is that it gives us the intuitive result that the disagreement between the two theories is not about what objects have final value or why, but rather the more abstract issue of how final value should be understood. Unfortunately, there is a worry that this intuitive result is won at the cost of emptiness and that relatively little has so far been said about what the concept of constitutive grounds really involves. As philosophers, we wish to better understand what theories like subjectivism are committed to when they claim that final value is determined by the presence of attitudes. Once the framework of constitutive grounds has been applied, most of what we learn is that constitutive grounds do not involve the relation between final value and its bearers or its makers. This is all well and good, but we may wonder what else the disagreement between subjectivism and objectivism could possibly be about: How ought we to interpret the role that is assigned to a factor when it is judged to serve as the constitutive grounds for final value? Is there a more positive story to tell about the commitments of theories like subjectivism and objectivism?

My aim is to look for answers to these questions. In so doing, I shall be trying to provide the framework of constitutive grounds with an interpretation that mirrors some of the claims made in the previous chapter. The interpretation therefore relies upon the same assumptions about the nature of final value: The first and most important is that final value can be understood in terms of the presence of reasons for attitudes. Final value can specifically be understood in terms of the existence of reasons to either favor or disfavor objects for their own sakes. The second assumption is that most normative properties, including the property of being a reason, must depend on other things. When an object has

\textsuperscript{192} The framework was originally developed by Rabinowicz in order to distinguish between different interpretations of preference utilitarianism (Rabinowicz & Österberg, 1996, pp. 1–2). According to what Rabinowicz and Österberg call the “satisfaction interpretation”, preference utilitarianism claims that the bearer of final value is the state of affairs that our preferences be satisfied. According to what they call the “object interpretation”, on the other hand, preference utilitarianism claims that the bearer of final value is the object of our preferences. It is in this latter interpretation that the concept of constitutive grounds gains its relevance. When they later move on to consider different objections to the object interpretation, this interpretation is revealed to rely on the plausibility of subjectivism as we here understand it (1996, p. 19).
value, or a fact is a reason, then there is something that affords them their normative status. In addition to these assumptions, the interpretation developed here finds inspiration in the view which has been defended by Schroeder (2007) and Sobel (2017) and others. This view states that the disagreement between subjectivism and objectivism should ultimately be understood as a disagreement about the factors that make facts into reasons. Here, the interpretation is specifically meant to give us a richer picture of what it means for something to serve as the constitutive grounds for final value. The analysis will be defended against several objections that are meant to show that it does not manage to capture everything that proponents of the framework of constitutive grounds want to say. Some of the objections can be easily answered, but others cannot.

Along the way, we shall let ourselves be seemingly sidetracked, as we look to the potential relevance of traditional views like *evaluative projectivism* for the proper understanding of the framework of constitutive grounds. In addition, we shall consider some of the more difficult challenges that subjectivism will be met with in its attempts to fully explicate final value and its associated phenomena. This includes the concept of value-makers which is presumed to fall under the domain of things that subjectivism is meant to explicate. Although these discussions will sometimes seem to be about the relative merits of subjectivism and objectivism as theories about the nature of final value, it is important to remember that the purpose of these discussions is not really to settle the disagreement between them. My hope, in other words, is not to cast either subjectivism or objectivism into doubt, as much as it is to highlight the many problems that may arise in the formulation of either theory. In the end, I just want to understand what we might commit ourselves to when we agree with subjectivism that final value is determined by the presence of attitudes, or when we agree with objectivism and deny that this is necessarily so.

Before we begin I wish to offer some further remarks about the apparent demandingness of subjectivism in this context. It has been pointed out that when we ascribe final value to an object, we are often unsure about precisely what makes it worthy of appreciation. However, without some sort of judgment as to what makes an object good or bad for its own sake, it may seem unreasonable to suppose that the properties of the object could figure in the intentional content of attitudes in the way that is here suggested. One response would be to claim that although we are often unable to consciously determine what makes certain things good or bad for their own sakes, there is still some such judgment lurking in the background when we judge things to have final value. Reconsider the case where we ask a friend to explain why a painting is beautiful. Upon hearing her inability to give an informative answer, we might
devise a test whereby we gradually change various aspects of the painting, like the composition or nuance of some of its colors. As a result, our friend might begin to think that the painting is horrendous.

Those who accept the line of response just mentioned will presumably claim that what has happened in this case is that, as a result of our changing the properties of the painting, our friend has become aware of some previously unconscious judgment on her part about what makes the painting beautiful. Although this is not a bad explanation, there is of course an alternative interpretation on offer. Instead of her becoming aware of a judgment which was previously lurking in the background, our friend may have acquired such a judgment. So, when the painting is being changed, she does not discover anything about herself, as much as she discovers something about the evaluative relevance of what is being changed. I am not sure how to go about settling which of these interpretations is right. I do not doubt that gradual changes like the ones we are imagining can make us aware of previously unconscious judgments, but I hesitate to say that this is always what happens in cases such as this. We want to be able to say that an object has final value without holding any judgment about what makes it so.

Another way of getting around the problem of over-intellectualization would be to maintain that subjectivism is only meant to apply to value and not to value-makers. What this means is that when it is applied to subjectivism, the framework of constitutive grounds should not make the second claim that was highlighted above. I am not entirely sure what to say about this suggestion, except that it seems to me to impoverish subjectivism and make it an incomplete theory. For this reason, I will here continue to assume that subjectivism, from the perspective of the framework of constitutive grounds, is committed to both claims highlighted above. Should this lead to problems, I will assume that they are problems specifically for subjectivism, rather than for the framework of constitutive grounds—or the interpretation of it that I am about to defend.\textsuperscript{193}

To reiterate, it is important to remember that I am myself not a proponent of subjectivism and so am not primarily concerned with trying to defend that theory here. Insofar as objections to subjectivism are considered in the following discussions, this is mainly so that they can help us highlight some of the features of the framework of constitutive grounds.

\textsuperscript{193} It is interesting to note that Fritzson (2014) at one point considers the suggestion that subjectivism should only make the second claim highlighted above and not the first. He attributes this suggestion to David Alm, who suggested it to him (in private communication). For more on the problem of intellectualization, see Fritzson (ibid, pp. 89–92).
Interpreting the Framework

According to the framework of constitutive grounds, subjectivism holds that attitudes serve as the constitutive grounds of final value, while objectivism denies that this is so. Rarely are we given a positive interpretation of what it means for attitudes to serve as the constitutive grounds of final value. Since constitutive grounds and background conditions both capture how things may turn properties into effective value-makers, there might be a suspicion that both are meant to do the very same job. We might even be tempted to regard them as involving the same concept and hold that they were formulated differently only because of figuring in the context of different discussions. Unfortunately, proponents of the framework of constitutive grounds have not done enough to deter us from this temptation so far, since they often tend to illustrate the role of attitudes for final value by using analogies and metaphors, the interpretation of which is itself a thorny issue. In the upcoming couple of sections, we shall take a closer look at some of those analogies and metaphors to squeeze some general insights out of them. My aim is then to provide a positive interpretation of the framework of constitutive grounds that avoids relying on analogies and metaphors, yet which is still capable of capturing most of the general insights.

In the first subsection to follow we shall be considering the metaphor involved in the theory of evaluative projectivism, which turns out to have a significant historical pedigree. According to the metaphor, value is in some sense projected, painted, or conferred upon the world by constitutive grounds. Subjectivism therefore claims that our having certain attitudes toward objects gilds them and endows them with that certain shine that we identify with the good and bad. We shall then briefly consider a more recent analogy involving the game of chess, which is meant to clarify the structural and modal features of constitute grounds. According to this analogy, things playing the role of constitutive grounds stay in the background of normative explanations rather than in their foregrounds. Although explanations that rely on these sorts of metaphors and analogies risk being misleading in many ways, I think there is enough of significance here that they nonetheless warrant closer attention. However, I shall endeavor to be relatively brief and to let others do the talking for me whenever it is possible. I thereby hope to avoid getting us completely bogged down in historical discussions and other interpretative issues that are only of a peripheral interest to the central topic at hand. That topic, to reiterate, concerns the most reasonable interpretation of the dependence claims made by the sorts of theories that I have already mentioned.
On Evaluative Projectivism

It is worth considering whether evaluative projectivism can help flesh out the framework of constitutive grounds a bit further, perhaps by elaborating on the first-person perspective with which subjectivism claims to be so intimately connected. At one point in the influential work *Leviathan* (1994/1651), Thomas Hobbes famously describes how objects in our surroundings produce in us certain processes or “motions” and that these, in turn, give rise to experiences or “fancies.” While doing so he urges that, though “the real and very object seem invested with the fancy it begets in us; yet still the object is one thing, the image or fancy is another” (ibid, section 1.4). The kind of fancies that Hobbes is here primarily concerned with happens to be color experiences. What he urges us to do, in other words, is to not conflate the colors of experience with something that is present in the objects that are experienced. The assumption seems to be that in the mind-independent world no color is really to be found. The source of all colors can in some sense be found within the minds of human beings. Hobbes later goes on to apply the same general line of reasoning to value:

> [W]hatsoever is the object of any man’s appetite or desire, that is it which he for his part calleth good: and the object of his hate and aversion, evil… For these words of good, evil, and contemptible, are ever used with relation to the person that useth them; there being nothing simply and absolutely so; nor any common rule of good and evil, to be taken from the nature of the objects themselves; but from the person of the man (ibid, section 6.7).

Color and value are similar, then, in that they are in some nebulous sense painted unto the world by people. They are projected into the world by beings with certain mental states. A century afterwards, David Hume similarly remarked that “the mind has a great propensity to spread itself on external objects, and to conjoin with them any internal impressions, which they occasion, and which always make their appearance at the same time that these objects discover themselves to the senses” (1978/1739, p. 167). When Hume

---

194 For more on the historical origins of projectivism, see Joyce (2009, p. 54). For more on the importance of first-order perspective for subjectivism, see Fritzson (2014, pp. 65–69).
195 For a much more thorough and careful interpretation of Hobbes and his brand of evaluative projectivism, see Darwall (2000).
196 For more on Hume and his brand of evaluative projectivism, see Olson (2014, pp. 21–42).
here speaks of the mind as spreading itself unto the world, imbuing objects with impressions that are internal to subjects, what he has in mind is not color but the phenomenon of causal necessity. He thought that we cannot literally experience the necessary connection between the events we judge to be causes and those that we judge to be effects. Instead, this causal necessity is attributed to events in the world based on observed regularities between them.\footnote{In so far as this tendency is not always under our conscious or voluntary control, it might seem very much like the tendency of people to ascribe value to objects that do not have them, but this similarity should not be exaggerated even by those who are skeptical of the existence of value. I would argue that there is a significant difference here irrespective of what we think of this ontological issue, for unlike causal necessity value \textit{is} part of our everyday experiences.} If one event is observed to follow another event enough times, observers will judge the latter to have been causally necessitated by the former. Causal necessity is in this way also sourced from the minds of human beings. A little later Hume follows Hobbes and goes on to apply the same general line of reasoning to value as well:

Thus the distinct boundaries and offices of reason and of taste are easily ascertained. The former conveys the knowledge of truth and falsehood: the latter gives the sentiment of beauty and deformity, vice and virtue. The one discovers objects as they really stand in nature, without addition or diminution: the other has a productive faculty, and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation (1978/1739, p. 167).

Causality and value are similar, then, in that they are inferred based upon the effects that certain experiences produce in observers. Other historical inspirations for evaluative projectivism can be cited, but Hobbes and Hume may be regarded as two of the main progenitors of the approach as it will here be understood.\footnote{Blackburn even refers to the approach as the ‘Humean picture’ of morality (1993a, p. 153).} The core idea behind their respective views on value, which of course differ in a lot of other respects, seems to be this: Value is not a part of the furniture of the world and whenever a value judgment is held or expressed, some attitudes tend to be or are necessarily present in the observer.\footnote{For criticism of the metaphor, see Brännmark (2002, pp. 65–69).} The attitudes of the observer are in some sense responsible for conferring value upon the world. Most prominent among contemporary defenders of this approach is arguably Blackburn. At one point he describes projectivism as the view that we have “sentiments and other reactions caused by natural features of things, and we
‘gild or stain’ the world by describing it as if it contained features answering to these sentiments, in the way that the niceness of an ice cream answers to the pleasure it gives us” (1993, p. 152).

This is perfectly in line with the views expressed by Hobbes and Hume. However, at another point Blackburn also states that we “project an attitude or habit or other commitment which is not nonnormative onto the world, when we speak and think as though there were a property of things which our sayings describe, which we can reason about, know about, be wrong about, and so on” (1984, pp. 170–171). He here amends the traditional evaluative projectivism of Hobbes and Hume with details about the ways we conduct our evaluative talk and practices. In this way, although the concept of constitutive grounds is not seen as involving a semantic relation as such, it may nonetheless have some semantic and conceptual elements to it. Specifically, the activity of spreading the mind upon external objects, by making these objects the target of attitudes, might be closely associated with certain ways of talking and thinking about final value. Finally, let us consider the following:

Let “Photoyou” refer to the picture of you and only you. Photoyou does depend on the properties of the photopaper (which I sometimes will refer to as ‘the paper’ or ‘the sheet’). Another paper with the same physical properties of the original photopaper would also put Photoyou on show. Now, some might object that I seem here to actually assume that Photoyou is a part of the photopaper, but surely, the objection goes, we must understand such an idea figuratively. Photoyou should not be located on the sheet of the paper; it is rather in the head of the person who looks at the photopaper. Whether or not this is the right ontological...

---

200 Blackburn has tried to explain why these ways of conducting our evaluative talk and practices are not necessarily mistaken. He refers to this explanatory program as quasi-realism (Blackburn, 1993a). He wants to show, in other words, that we may continue to talk of things like the truth of evaluative statements and the mind-independence of value and all the rest of it, without committing ourselves to there actually being any such things. Other contemporary supporters of the approach are not quite as unwilling to accept that our evaluative talk and practices are essentially mistaken, among them being Mackie (1977) and Joyce (2009). This difference in outlook is of course due to a metaethical disagreement between Blackburn on the one hand and Mackie and Joyce on the other. The former is a proponent of expressivism and thus claims that the primary purpose of evaluative statements is to express the speaker’s preferences, attitudes, and desires (Blackburn, 1993a). The latter, as we have seen earlier, are proponents of cognitivism and insist that our evaluative statements are mainly meant to describe some aspect of the world that turns out not to exist (Joyce, 2001).
position to take with regard to Photoyou (and I am not arguing
for any position here), the illustrative force of this example does in
fact very much derive from the belief that Photoyou is not
necessarily on the paper but in someone’s head. The ground (in
this case the constitutive ground) of Photoyou is to be looked for
in the fact that the observer perceives the properties of the paper;
Photoyou is, the analogy goes, constituted by your perspectival
look on the paper (the supervenience base of Photoyou). If you are
a projectivist of some kind, you will find reason (most probably in
your phenomenological experience) to say that Photoyou is also
“out there,” put there by you (Rønnow-Rasmussen, 2003, p. 253).

Aside from elaborating on the projection metaphor, and giving it a more
modern twist, Rønnow-Rasmussen here seems to suggest that there are at least
two understandings of evaluative projectivism, one of which is more realistic
than the other. I think this is right, but my suspicion is that both
understandings, if applied to subjectivism, risk undermining the disagreement
that it has with objectivism.201 This is worth exploring before we move on. In
the face of evaluative projectivism, subjectivism seems to become a theory about
what happens in certain parts of our minds and how things are represented
therein when we either favor or disfavor objects for their own sakes.202 Earlier, in
the introduction to this chapter, I made things easy for us by describing
objectivism simply as the negation of subjectivism. However, it is not
implausible to suggest that objectivism need not take any stance on what
happens in certain parts of our minds and how things are represented therein

---

201 This would be reminiscent of what has happened with traditional metaethical disagreements
between non-cognitivism and cognitivist realism. In the good old days, the disagreement was
to a large part about whether there are mind-independent evaluative facts and whether
evaluative statements can possibly be true or false. While non-cognitivists gave a negative
answer to such questions, cognitivist realists gave a positive one. However, to accommodate
our evaluative talk and practices, a new generation of non-cognitivists began to make
compromises (e.g. Blackburn, 1984, 1993a; Gibbard, 1990, 2003). By relying on a
deflationary understanding of what it means to say that something is true, false, or a mind-
-independent fact, these non-cognitivists tried to show that such ways of speaking are fully
compatible with their views. In other words, their point was that we can continue to talk of
truth, falsehood, and mind-independent facts, even if turns out that evaluative statements have
a primarily expressive, commissive, or directive illocutionary character. I would suggest that
despite the appearance of progress that comes with this conciliatory approach, the
compromises of non-cognitivism merely obscure a divide which remains very deep and real.

202 Rabinowicz has admitted (in private conversation) that he got the term from phenomenology,
where constitution is often understood as a creative process on the part of the mind.
when we favor or disfavor objects for their own sakes. If this is right, then there is a temptation to suggest that while subjectivism is mainly focused on explaining the nature of evaluation, objectivism is not.\(^{203}\) This means that the two theories risk talking past one another. When subjectivism subscribes to the claim that final value is determined by attitudes and objectivism denies that this is so, they are interpreting the claim in separate ways. Subjectivism understands it in broadly psychological and linguistic terms, but objectivism does not. Perhaps objectivism understands it in more explicitly metaphysical terms, meaning that it is the only theory that is really about the nature of final value.

Proponents of subjectivism tend to claim that final value is not a genuine property or relation in the world. This claim is obviously about the nature of final value and easy to understand. However, the fear is that whatever else they may say about final value ends up not being about its nature as much as about what happens in certain parts of our minds and how things are represented therein when we either favor or disfavor objects for their own sakes. One way of retaining the disagreement and granting subjectivism the claim to also be about the nature of final value is to interpret it a bit more literally than we have done so far. The theory of subjectivism would then be understood as the suggestion that final value may be a genuine property or relation in the world after all, but to the extent that final value exists, it does so only in a subjective sense.\(^{204}\) What this means is that it is in the nature of final value, conceived as a genuine part of the furniture of the world, that it can be observed only from the first-person perspective and never from the third. In so far as this literal interpretation of evaluative projectivism even makes sense, it seems as if it would also undermine the disagreement between subjectivism and objectivism far too much. For proponents of the two theories would then effectively agree that final value exists but would disagree about what sort of existence final value has. Whereas one theory claims that final value exists only subjectively, the other claims that final value exists objectively. This is an ugly result which we should probably avoid. When subjectivism claims that final value is conferred upon the world by the

\(^{203}\) For more on projectivism as a psychological thesis, see Joyce (2009) and Pölzler (2016).

\(^{204}\) Joyce once said: “Projecting attitudes is not like projectile vomit” (2009, p. 60). According to the view just mentioned, it may be more like projectile vomit than we previously supposed. Compare also with Brännmark, who states that proponents of constitutive grounds borrow terminology which “makes it sound as if these writers can grant values an air of ontological respectability that is not really backed up by the ontology at their disposal” (2002, p. 66). This charge is very reminiscent of Quine, when he, in a more general discussion, accuses some people of engaging in “philosophical double talk, which would repudiate an ontology while simultaneously enjoying its benefits” (1960, p. 242).
attitudes of observers, it should not be taken quite this literally. We need some other way, then, of reconciling evaluative projectivism with the strong disagreement between subjectivism and objectivism—or give up the metaphor.

It would be of no help to claim here that subjectivism is a theory about the nature of final value in much the same way that certain forms of idealism are theories about the nature of the world. The idea behind such a response would be that some forms of idealism might be regarded simply as very general versions of subjectivism. Rather than saying that final value is constituted by mental states, idealism insists that, as a matter of fact, everything is mental in nature. If it is reasonable to regard idealism as a story about the nature of the world, then it must also be reasonable to regard subjectivism as a story about the nature of final value. There is much to be said about this analogy. In fact, this is a juncture at which we can easily get lost in what, in the current context, must be regarded as rather peripheral historical and interpretive issues. I shall therefore try to keep matters as short as I possibly can. My suspicion is that there are significant differences between subjectivism and idealism that undermine the analogy just made. They concern the issue of realism. As I have said, proponents of subjectivism tend to deny the existence of what they claim is constituted by attitudes, but proponents of idealism do not. Proponents of idealism do not deny the existence of the world but are only skeptical of the way its existence is conceived by common-sense understanding. The world exists, according to idealism, but it just so happens that the world is not what we tend to think.

The main point I wish to make here is perhaps that the projection metaphor seems to give us some additional clues regarding the understanding of constitutive grounds and the disagreement between subjectivism and objectivism. However, the projection metaphor invites certain interpretive problems that proponents of the two theories would do well to steer clear of. Regardless, it is also doubtful whether the projection metaphor is enough to fill in the gaps that were highlighted at the start of this discussion. We are still in need of a positive interpretation of the framework of constitutive grounds that clarifies the disagreement and ensures that proponents of the respective theories are not just talking past one another. Preferably, the interpretation should not rely on further analogies and metaphors which are, on their own, very difficult to interpret, and which invite the kinds of discussions that I have just been hinting at. However, I will return to the alleged importance of the first-person perspective later in my discussions, when we will consider some problems with the interpretation that I am about to defend.
On the Chess Analogy

The previous section dealt with a metaphor that appears to emphasize a psychological component inherent in the framework of constitutive grounds. This section will be mainly concerned with the structural and modal features of the framework of constitutive grounds. The framework of constitutive grounds already appears to be mostly about structure and modality, as its primary purpose is to carve out conceptual space for a new type of relation. To reiterate, this relation is meant to explain how things like attitudes can play a significant role for final value in whatever world it is to be found, but without being among the factors that final value accrues to or the factors in virtue of which it accrues. We shall give these aspects an even closer look by considering an illustrative analogy. At one point in their early discussions of the framework of constitutive grounds, Rabinowicz & Rønnow-Rasmussen attempt to clarify its basic points by making comparisons to the game of chess and its rules:

That a certain move in chess is admissible is a feature that supervenes on the internal properties of the move and of the situation on the chess-board. But the constitutive grounds of its being admissible are to be found in something external—in our conventions that determine the game of chess. Similarly, according to some preferentialist conceptions of value, preferences or attitudes may bestow a value on the object toward which they are directed. Still, if the object is being preferred for the features that are internal to it, then this externally constituted value is intrinsic: it supervenes on the internal properties of the object, precisely those properties for which the object is being preferred (2000, p. 37).

Now, it is important to stress what should already be obvious at this stage that constitutive grounds are not to be understood in terms of parthood. The claim that attitudes serve as the constitutive grounds of final value is not the same kind of claim that a piece of mud is part of a statue, that agency is a part of personhood, or that a mixture of rum and coke is a part of a Cuba Libre. The chess analogy is therefore potentially a bit misleading, since it is unclear whether the conventions determining the game of chess are not even partly constitutive of the game itself in this mereological sense. Nevertheless, the analogy is still useful for gleaning some general observations about the behavior of constitutive grounds. We will consider a few of those observations here as they are relevant
to the more constructive discussions to follow. Proponents of subjectivism claim that favoring close friendships for their own sakes determines that close friendships have final value, even when no final value accrues to the attitudes themselves. Nor is the favoring of close friendships for their own sakes responsible for making knowledge have final value in the narrow sense that involves evaluative resultance. So far, then, everything is in order.

To explain what makes close friendships good for their own sakes we would point to the properties that close friendships tend to have. We might say that close friendships are good for their own sakes in virtue of their extrinsic property of making our lives easier, or in virtue of their intrinsic property of being characterized by a mutual respect and felt concern. These are the properties for which close friendships are favored for their own sakes. Intuitively, once we go beyond this story and elaborate on the broader role played by attitudes for this final value, we have started to venture into a more abstract area of value theory. The same point seems to apply to the chess analogy: When considering the soundness of a chess move, the consequences of the move on the board will figure in the foreground of our minds, while the broader norms determining the game of chess are relegated to the background or kept entirely from view. When asked to explain why we have made a certain chess move, moreover, we will formulate the explanation in terms of the consequences of the move on the board. And when we go beyond this and elaborate on the broader norms determining the game of chess, we are not considering questions that are asked within the context of a specific game of chess, as much as we are answering questions about the game of chess itself.

The chess analogy also invites observations about the modal character of constitutive grounds, some of which can be brought out by considering the common variability objection to the theory of subjectivism. There is a worry that when we consider a possible world featuring a different distribution of attitudes, we will be forced to conclude that, in that possible world, there will also be a different distribution of final value. To many, this might seem like an absurd consequence that casts doubt on the validity of subjectivism as a theory about the nature of final value. Proponents of subjectivism have responded to the variability objection by allowing for actual attitudes to serve as the constitutive grounds of final value across different possible worlds (Rabinowicz & Österberg, 1996, pp. 19–23; Fritzson, 2014, pp. 71–80; cf. Rønnow-Rasmussen, 2003, p. 258). In other words, then, that knowledge is favored for its own sake in the actual world ensures that knowledge has final value in whatever possible world it
is to be found. Rabinowicz and Rønnow-Rasmussen elaborate on this point as follows:

If our preferential attitudes are seen as the constitutive grounds of an object’s final value, then it is natural to suppose that they confer value on the object not only in the actual world but also in other possible worlds toward which these attitudes may be directed. Consequently, they might confer value on objects even in those worlds in which our attitudes toward these objects would have been different from what they actually are. Thus, unlike the dependence base, the constitutive grounds need not be present in all those possible worlds in which the object is supposed to be valuable (2000, p. 37).

A similar observation can probably be made by using the chess analogy, if we are allowed enough leeway. There is a possible world where people like to play a game that they refer to as ‘chess’, but which is determined by conventions that are somewhat different from those that determine the game we all know as ‘chess’. It seems clear that even though we both use the term ‘chess’, the games we are talking about when we use that term are potentially very different from one another. This depends on how similar the relevant conventions are that determine the games we are playing. If they are different enough it might even make sense to say, given our specific perspective, that the game of chess has never actually been played in that possible world. This means that constitutive grounds could be taken to have a modal character of a very rigid nature, since their influence stretches beyond the limits of the specific possible worlds where they are found. We should stay on this topic, for it has the potential of being

---

205 There are cases where actual attitudes are sensitive to beliefs about hypothetical attitudes. For example, we may not have a positive attitude to an object situated in another possible world if we lack any positive attitudes to the object in that possible world (Fritzson, 2014, p. 49).

206 This sort of maneuver is reminiscent of what proponents of two-dimensional semantics have been saying of rigid designators and *aposteriori* necessities (e.g. Chalmers, 1996, pp. 56–65; Jackson, 1998, pp. 47–52). Suppose there is a possible world where everyone uses the term ‘water’ to speak of the liquid XYZ. Does that mean that water is XYZ in that world? According to proponents of two-dimensional semantics, the answer is ‘yes’, as long as we center that possible world and treat it as if it were the actual one. From the perspective of our world, however, the answer is in the negative: In fact, because the term ‘water’ refers to H2O in our world, it may even make sense for us ignore what its inhabitants might say and deny that there is in fact any water in that other possible world. Despite their similarities, it should be mentioned that the response offered by subjectivism to the problem does not necessarily rely
crucial to our understanding of both the concept of constitutive grounds and the theory of subjectivism.

It has been pointed out that an object can retain its value if the old resulstance base is replaced by a new one, which grounds the same value. This depends, among other things, on whether there is something else in virtue of which the object could end up being good or bad for its own sake. Once an arrangement of factors is changed, whatever arrangement is left may still be viable as a resulstance base for final value. For example, reshaping a beautiful sculpture may result in a qualitatively very different yet equally beautiful sculpture. Something similar could be said of background conditions, which, as we have seen, do not even have to be parts of complete explanations for final value to begin with. Whether the same is true for constitutive grounds arguably depends upon how we look at other possible worlds. Suppose the attitudes toward a given object serves as the constitutive grounds for the final value of that object. Imagine also that we consider a possible world where these attitudes do not exist but do so from the point of view of that world. This means that we center that possible world as the place from which the value judgments are to be assessed. From this perspective the object in question does not have final value.

Proponents of subjectivism are here likely to rely upon a kind of evaluative solipsism which ties our value judgments to a first-person perspective. According to subjectivism, to judge anything to be good or bad for its own sake is to experience it from the first person and this perspective will inescapably involve the presence of actual attitudes. When we start talking in the technical terms described earlier and imagine what it would be like if other attitudes were actual, then we start looking at things from the third person, which is really a perspective from which nothing can be said to have final value.\textsuperscript{207} Even in light of this, however, there is a small remnant of the original worry which can be formulated without the use of possible worlds in this way. The objection would then try to capture the intuition that, even if our attitudes are different than what we currently think they are, then the same things would still have final

---

\textsuperscript{207} Issues of perspective turn out to be important to subjectivism and are sometimes used on their own to characterize its disagreement with objectivism. According to the former, final value can in principle only be seen from the perspective of those with attitudes. From the point of view of the universe, on the other hand, nothing really matters. Of course, subjectivism enables us to establish, from the third person perspective, that a person has certain attitudes for an object and that it will therefore seem valuable to him. But this is a different matter.
value. This is the objection that Huemer has in mind when he stresses that the relevant circumstance should be imagined in the indicative way:

One way of countering this objection would be to ‘rigidify’ evaluative terms—that is, one might propose that ‘good’ always has for its extension what we actually approve of, even when we are discussing counter-factual situations in which we wouldn’t approve of those things. Thus, the reason we don’t have to say that if everybody approved of Hitler then Hitler would be good, is that we presently disapprove of Hitler, and we continue to disapprove of him even when we’re thinking about that counter-factual situation. One consequence that can not be escaped by this _ad hoc_ maneuver, however, is that any sentence of the form, ‘If we approve of x, then x is good,’ where the conditional is indicative (not subjunctive), must be true (1996, section 4).

It is left to the reader to decide how damning this kind of objection against subjectivism really is. If it is true that quite a lot of our attitudes are different from what we often think they are—or what we have, for the sake of civility and other factors, convinced ourselves that they are—then perhaps some of us will agree that the objection holds some water.\(^{208}\) Regardless, the observations made so far have taught us something important about the behavior of constitutive grounds. Chief among them is that it is at least meant to be possible for constitutive grounds to exert their influence across different possible worlds. When one thing serves as the constitutive ground of another, the objects in question may in fact end up being on opposite sides of a modal border. Another lesson is that, according to subjectivism, constitutive grounds are in some sense

\[^{208}\text{If there is something to the objection, then I also think it retains its force even if proponents of subjectivism were to shift their focus to relative value. This response was suggested to me (in private conversation) by Fritz-Anton Fritzson. Proponents of subjectivism may well argue that theirs is ultimately a theory about value relative to a person, since there is no other kind of value. It is not clear what this kind of value is meant to be, but what is clear is that it is not the kind of value attributed to something when we say that it is good for a person. The idea, at any rate, is that what Huemer’s original variability objection manages to show is merely that relative to the subjects with other attitudes, other things will be valuable. Admittedly there is nothing particularly strange about this. However, I do not think that this is a response at all, for the problem would then be that what is valuable relative to us may be different than what we think it is. It is contrary to our intuitions that this difference should depend upon what attitudes we actually end up having. The shift in focus to relative value does not afford subjectivism with anything that is not already present in the account above.}\]
bound up to a particular first-person perspective. Though this latter point is obscure, it can perhaps be used to correct an assumption that seems to be part of the last modified version of the objection formulated by Huemer in the quoted paragraph. On his understanding subjectivism rigidifies evaluative terms.

The idea that Huemer has in mind is presumably that locutions like ‘good for its own sake’ could be likened to the rigid designators of Saul Kripke (1972), picking out objects in other possible worlds in accordance with the attitudes that are actually held. This interpretation was also invited by my earlier attempts to compare the case of final value to the case involving chess. There is arguably something very problematic about this idea. A person who opposes the theory of subjectivism or denies that only actual attitudes serve as the constitutive grounds for final value is clearly not confused about the semantic features of value terms. To suggest otherwise would be to stack the deck against objectivism or any other alternative which implies the falsity of subjectivism. Indeed, Rabinowicz & Österberg themselves explicitly state that when attitudes are said to serve as the constitutive grounds for the good and bad, “this is a claim about the nature of value and not about the meaning of value statements” (1996, p. 26). Rather than rigidifying evaluative terms, then, we should take subjectivism to rigidify attitudes in some sense. As we engage in evaluative talk and practices, we confer final value upon things not only in the actual world, but in other possible worlds as well. Spelling this out in a way that does not end up undermining the disagreement between subjectivism and objectivism turns out to be surprisingly difficult.

Toward a Hyperintensional Interpretation

All proponents of subjectivism must agree that without the presence of attitudes, there would be no value, and nothing would matter. There is therefore a wide sense in which they all must claim that the presence of attitudes makes facts reasons. Sobel (2017), who is among the most ardent supporters of subjectivism in the past few decades, writes that a fact “is a practical reason for one to do something due to the existence of some sort of not truth-assessable favoring attitude one has that would be served by doing so” (ibid, p. 17: my emphasis). He also states that facts “are reasons for a person if and only if and because of what she values and favors” (ibid, p. 17: my emphasis). Sobel seems to me to be using the emphasized terms in a mostly generic sense here. There remains some unclarity, then, as to how he understands the sense of determination that is
appealed to by subjectivism. Sobel perhaps comes closest to a technical explication like the one I shall be defending when he later declares that subjectivist theories are “best understood as accounts of the truth-maker in a certain domain” (ibid, p. 117: my emphasis). I take it that truth-making is a relation between propositional entities, and so the idea advanced by Sobel seems to be that facts about attitudes serve as the truth-makers for facts about what we have reasons to do. Although my concern is not with truth-makers, this seems to capture roughly the same idea as the interpretation I will be defending.

Schroeder (2007) is another prominent supporter of subjectivism. He states that subjective reasons for a person to do something are the kinds of reasons whose presence “depends on some feature of that person’s psychology, such as what that person desires, wants, likes, or cares about” (ibid, p. 1: my emphasis). He then states that however the relevant sense of explanation works, “all reasons are explained by psychological states in this same way” (ibid, p. 2). Schroeder gets more technical when he tries to elaborate on these points by appealing to the thesis of hypotheticalism: “For R to be a reason for X to do A is for there to be some p such that x has a desire whose object is p, and the truth of R is part of what explains why X’s doing A promotes P” (ibid, p. 57). When taken together, this selection of quotations expresses the idea that the presence of reasons is explained by the obtaining of the facts that constitute them, together with the obtaining of facts about what agents happen to favor. Curiously, Schroeder at one point formulates his views in terms of background conditions as well:

In general, according to Hypotheticalism, the objective normative reasons for x to do A are the things which help to explain why X’s doing A promotes P, where P is the object of one of X’s desires. And the fact that R helps explain why X’s doing A promotes P isn’t part of the reason—only R is the reason. And R needn’t mention X at all. So there is a distinction between reasons and background condition (ibid. p. 29).

Schroeder seems to be saying that facts about attitudes that help explain the presence of reasons should be regarded as background conditions. They are background conditions, he says, because facts about the attitudes are not themselves part of the relevant reasons. I certainly agree that subjectivism should not be understood as saying that facts about attitudes must be parts of reasons. This goes plainly against the aims of the framework of constitutive grounds. And since Schroeder sees the major point of contrast as being background conditions and the factors that form parts of reasons, his understanding of background
conditions seems to mirror the one that I defended in the last chapter. According to Schroeder, subjectivism is a theory about attitudes being parts of the explanation for the presence of reasons in some sense. Whether that sense is ultimately identical to the one I shall have in mind is a bit uncertain. All I can do for now, then, is hope that the overlap between my interpretation and the views defended by many proponents of subjectivism, like Sobel and Schroeder, is bigger than appearances suggest.

Many will agree that close friendships are good for their own sakes. Suppose the value of close friendships results from the extrinsic property of their parties having a history of caring for one another’s welfare. The fact that close friendships have this extrinsic property is the reason to favor close friendships for their own sakes. The fact that close friendships have this extrinsic property must be a reason in virtue of, or as a result of, something else. The theory of subjectivism claims that what affords the fact in question with the normative property of being a reason is the presence of attitudes. More specifically, the role of constitutive grounds is played by the favoring of close friendships for their own sakes. The fact that close friendships have the extrinsic property ends up being a reason precisely because the extrinsic property is among the features for which close friendships are favored for their own sakes. This explains why the fact is treated as a value-maker in this context.

Suppose next that the value of close friendships results from some feature that is intrinsic to them, like, say, the property of their involving a mutual sense of love or respect. The fact that close friendships have this property is the reason to favor close friendships for their own sakes. Once again, the fact that close friendships have this intrinsic property must be a reason in virtue of, or as a result of, something else. The theory of subjectivism claims that what affords the fact in question with the normative property of being a reason is the presence of attitudes. More specifically, the role of constitutive grounds is played by the favoring of close friendships for their own sakes. The fact that friendships have the intrinsic property just mentioned ends up being a reason precisely because the intrinsic property is among the features for which close friendships are favored for their own sakes. Objectivism rejects this and maintains that features not involving attitudes can explain how facts such as these come by their normative status.

If the interpretation just presented is correct, then it is not difficult to see why subjectivism is not committed to the claim that attitudes, or facts about attitudes, are themselves good or bad for their own sakes. Furthermore, it should not be difficult to see why subjectivism is not committed to the claim that attitudes, or facts about attitudes, make objects good or bad for their own sakes.
What attitudes do is endow certain facts with the normative property of being a reason. Here, the framework of constitutive grounds is applied to subjectivism again, this time with its interpretation built in:

**Value Constitution Elaborated:** Final value accrues to objects only to the extent that the objects are favored or disfavored for their own sakes, because attitudes are parts of the resultant base explaining why facts have the normative property of being reasons to favor or disfavor the objects in this way. This clarifies why attitudes themselves are not bearers of final value, unless they too are among the things that are targeted by attitudes in this way.

**Maker Constitution Elaborated:** Final value accrues to objects in virtue of certain factors only to the extent that the objects are targeted by attitudes in which the factors in question are represented as value-makers. The reason is that the attitudes in which the factors are represented as value-makers are parts of the explanation why there are reasons to favor or disfavor the objects for their own sakes to begin with.

We want to understand the commitments of theories like subjectivism, which claims that final value is determined by the presence of attitudes. According to the interpretation that has just been put forward, this should be understood along the lines of the suggestions made in the previous chapter, in terms of attitudes serving as the base for the normative property of being a reason. Although the exact nature of the relation remains a problematic issue, relying upon it in this context still brings with it a theoretical gain. For in addition to the negative characterization given to us by the framework of constitutive grounds, we now have a positive story to tell about how final value can be determined by the presence of attitudes. This story entails that the determination spoken of by subjectivism is no longer treated as a mere placeholder, ensuring that subjectivism can remain neutral about what has final value or what makes things good or bad for their own sakes. The story explains why theories like subjectivism can remain neutral in this sense.

This benefit can also be highlighted another way, by noting that philosophers have occasionally feared that the truth of subjectivism would be incompatible with the possibility of intrinsic final value (e.g. Moore, 1922: 269; Korsgaard, 1983: 174). If final value is determined by the presence of attitudes, then it must be instantiated in virtue of, or as a result of, the presence of attitudes as well. As
a result, only attitudes themselves can have intrinsic final value, and everything else must have extrinsic final value. The framework of constitutive grounds allows subjectivism and objectivism to be neutral about what makes objects good or bad for their own sakes. This benefit is preserved by the interpretation just put forward. My interpretation commits subjectivism and objectivism to a story about what makes facts reasons for attitudes, not what has final value or what makes objects good or bad for their own sakes. This leaves the door wide open for the concept of extrinsic as well as intrinsic final value. I will defend this point further in just a moment.

An initial worry is that the interpretation just defended does not manage to avoid understanding subjectivism and objectivism as theories about what makes objects good or bad for their own sakes. A version of this objection was already mentioned in the previous chapter, but perhaps it is worth reconsideration within this narrower context. Suppose an object has final value in virtue of, or as a result of, there being certain facts present, and these facts are relevant inasmuch as they are afforded their normative status by attitudes. It follows that there is a general sense in which the object has final value in virtue of, or as a result of, the presence of said attitudes as well. To reiterate, the mistake here is to forget that value-makers should be identified with the things we would first point to when asked to explain why an object has final value. These are the facts that constitute reasons for attitudes. The explanatory bases figuring at a lower level, which account for why facts are reasons for attitudes to begin with, do not constitute value-makers.

Proponents of subjectivism might be reluctant to accept that there are reasons for attitudes on account of this jeopardizing the priority of attitudes over the normative. It seems to me that this is not any part of the interpretation just laid out. If an object is not the target of any attitudes, then there are no reasons to favor an object for its own sake. Only once the object becomes the target of attitudes are we given reasons to favor it for its own sake. This means that attitudes are prior to the normative after all and so proponents of subjectivism need not worry on this account. This commits subjectivism to a form of bootstrapping, whereby favoring something gives rise to reasons for favoring, but this seems to me to an old problem. Indeed, it is difficult to imagine how we could avoid this result and still have attitudes be prior to the normative. Proponents of subjectivism can accept the bootstrapping result or restrict the set of things that we can have reasons to do.

Constitutive grounds were described earlier as having a very rigid character in that they have an influence that stretches across possible worlds. The question is how compatible this observation is with the interpretation so far defended. The
intuition is that when the normative property of being a reason is instantiated, then this cannot be in virtue of, or as a result of, features that are not situated within that world. There is therefore a discrepancy between the relation I have in mind and the one that is hinted at by the framework of constitutive grounds, which means that one cannot be reduced to the other. One stubborn line of response would be to insist that the normative property of being a reason can be instantiated in virtue of, or as a result of, factors that it is modally separated from. According to the interpretation just laid out, subjectivism commits itself to stating that facts are afforded the normative property of being a reason by the presence of attitudes. It seems to me to make sense to take this story further and suppose that a proponent of subjectivism could maintain that attitudes can give facts their normative status in whatever worlds the facts are found.

There is another way of trying to avoid the counterintuitive effects of allowing a normative property and its resultance base to be situated on opposite sides of a modal border. However, this way requires us to pull off some strange maneuvers whereby we describe possible resultance bases in terms of what is actually the case. Suppose again that there is a possible world where we have other attitudes than we actually have. Our actual attitudes are directed toward many objects in that possible world which are not favored by its inhabitants. In the actual world people tend to like chocolate ice-cream, let us say, but none of the inhabitants of the possible world like chocolate ice-cream. Now, it might be understood as a fact within that possible world that the objects in question are the targets of our actual attitudes. If this is right, then perhaps this fact could also be part of the resultance base for the normative properties within that world. This would mean that no modal border would have to be crossed.

209 Here it is perhaps relevant that subjectivism is often formulated as a theory not about the relevance of the attitudes we actually have, but of the attitudes that we would have in idealized circumstances (e.g. Sobel, 2017). It might be thought that idealization of this kind will unavoidably rely upon intuitions about what has value irrespective of what attitudes we might have. This would mean that there is something profoundly objectivist in the entire endeavor of idealization. For this reason, it is often emphasized that in so far as subjectivism ought to speak of attitudes that we would have in idealized circumstances, the relevant form of idealization should be understood in procedural terms. What this means is that subjectivism does not rely directly upon intuitions about what has value and what attitudes we should therefore have, but rather on the methods and processes by which we adopt attitudes in the first place. Now, regardless of whether we understood idealization in substantive or procedural terms, there is an obvious temptation to understand idealized attitudes as simply being hypothetical ones. See also Risberg (2017) for an excellent critique of idealization-accounts of value.

210 Gratitude is owed to Wlodek Rabinowicz who (in private communication) pointed out to the possibility of this line of response, although, it should be mentioned, he did not endorse it.
Let us recall the demarcation mentioned briefly in the last section, on which the background conditions for final value are the factors that provides facts with the normative property of being a reason without themselves being parts of the reasons. For this entails that the interpretation we have in mind does not necessarily result in subjectivism and objectivism being theories about background conditions. At the very least, this is not all they are. Suppose that a friend is sad and that this is a reason for us to comfort her. It seems plausible to suggest that this is a reason in virtue of, or as a result of, the sadness of our friend. In other words, her property of being sad affords the fact that she is sad its normative status. Because her sadness is part of the reason we have to comfort our friend, it is not a background condition. Now, given the discussions in previous paragraphs, we might want to add an additional requirement by stating that although constitutive grounds and background conditions involve the same phenomenon at bottom, background conditions are the specific reason-makers that do not form part of the reasons themselves, and whose influence is confined within possible worlds.

The last objection I will be considering is about the neutrality which is meant to characterize the disagreement between subjectivism and objectivism. There are two different issues here which may cause concern, one conceptual and one substantive. Even those who agree that value is reducible to the presence of reasons might feel uncomfortable with its use in the present context. The idea is that to understand the disagreement between subjectivism and objectivism, we should not have to invoke conceptual views of this kind. More generally, the intuition is that subjectivism and objectivism are meant to be entirely neutral about how we should understand the connection between value and reasons. The interpretation we have laid out does not manage to accommodate this intuition. A similar worry crops up from a more substantive point of view. The interpretation we have laid out places the disagreement between subjectivism and objectivism outside one substantive area of enquiry where it does not belong. It allows for proponents of the two theories to agree on what has value and what makes objects have value. The disagreement is instead taken to be about what makes facts reasons. The objection is that a plausible interpretation should also allow proponents of subjectivism and objectivism to agree on this issue. More generally, the intuition is that subjectivism and objectivism are

---

211 This kind of example is just meant to illustrate the general demarcations on background conditions which was mentioned in the earlier chapter. Subjectivism would of course deny that the fact that our friend is sad is part of what gives us a reason to comfort her. This role is rather played by the fact that we not like her being sad, or something similar.
meant to be outside all substantive areas of enquiry. For example, Rønnow-Rasmussen writes:

Therefore, the first thing to keep in mind is that the distinction between subjectivism and objectivism is not a distinction between two different substantive views. The issue between subjectivists and objectivists is (if we confine ourselves a moment to the ethical aspects of the evaluative) metaethical; advocates of these theories give us different answers to how we should best understand and analyze value claims. However, the positions in themselves do not force us to take any evaluative stand. In other words, nothing in principle prevents subjectivists and objectivists from endorsing the same evaluative judgments (2003, p. 248).

Theories cannot be informative and yet remain neutral about all substantive and conceptual issues, but neutrality about the specific issues just mentioned would still be nice. When all else is equal, it is of course true that the fewer controversial assumptions theories have to manage the better.212 Of course, that the interpretation I have defended fails to be neutral about some issues could be regarded as a disadvantage either of the interpretation or of the framework of constitutive grounds. If the interpretation is discarded, a better alternative should be presented in its place, lest neutrality be won at the cost of informativity once again. The interpretation developed here entails that when proponents of subjectivism are asked to explain the relation between final value and the presence of attitudes, they can do more than list all the ways in which the relation is not to be understood. It might be true that the price of these benefits is too high. With the exception of Schroeder, I expect that most will insist that the interpretation is not in line with the spirit of either subjectivism or objectivism. My hope is that the preceding discussion will encourage these philosophers to say something more constructive about constitutive grounds.

I have tried to show that distinguishing reasons from the normative property of being a reason invites innovative ways of understanding some old and controversial concepts. Among them is the concept of background conditions, which was explained in the previous chapter in terms of the factors that make facts reasons for attitudes. I argued that this explanation has two significant benefits: The first is that it manages to capture the substantive intuitions which

212 See also Rønnow-Rasmussen (2003, pp. 251–253).
inspired the need for the concept of background conditions in the first place. This is the intuition that there are factors that contribute to the presence of value and reasons, without being parts of value-bearers, value-makers, or reasons themselves. The second benefit is that it manages to be theoretically parsimonious in the sense that it does not take background conditions to involve an entirely new kind of relation the likes of which we were not already committed to. Next up were traditional views on the nature of normativity, specifically subjectivism and objectivism, which were similarly interpreted as making claims about the things in virtue of, or as a result of, which facts are endowed their normative status. Whether that interpretation does justice to the framework of constitutive grounds is a controversial matter. If there is an alternative interpretation which manages to do more justice to the framework of constitutive grounds, but which is also more capable of avoiding some of the problems that have just been mentioned, I would welcome it with open arms.

For example, I concede that evaluative projectivism may have been dismissed a bit too quickly in the preceding discussion and that the key to interpreting the framework of constitutive grounds is still to be found in ideas about the centrality of the first-person perspective. Perhaps it is here, and in the deeper discussions about the connection between subjectivism and idealism, that the conceptual and substantive neutrality which is apparently missing from my own account can be won. I do not rule this out and would acknowledge that my own difficulty in understanding evaluative projectivism and its emphasis on the first-person perspective might have less to do with its obscurity and more to do with philosophical bias and my unfamiliarity with the philosophical tradition in which it the theory is entrenched, including phenomenology. However, from my own first-person perspective, the greatest difficulty for this philosophical tradition seems to be that it makes subjectivism out to be too much of a psychological theory and that it risks significantly undermining the disagreement that its proponents are generally thought to have with defenders of objectivism. Perhaps this risk is worth the cost of retaining the conceptual and substantive neutrality which is also thought by adherents of the framework of constitutive grounds to characterize the disagreement between the two theories.

Concluding Remarks

This and the previous chapter could be regarded as companion pieces. They were dedicated to understanding that part of the dependence of final value
which does not involve evaluative resultance. When objects are good or bad for their own sakes, this may have something to do with factors to which we would not want to assign the role of value-makers. The previous chapter dealt with the concept of background conditions and in the present chapter I have endeavored to become clearer about the framework of constitutive grounds. This framework was developed to make better sense of two traditional theories about the nature of final value, namely subjectivism and objectivism. Subjectivism is the theory which claims that final value is in some sense determined by attitudes, without attitudes themselves being value-makers or the things that final value needs to accrue to. Objectivism denies that final value must be determined by attitudes in this sense. I made the argument that although the framework of constitutive grounds may seem intuitively plausible, this might be a result of its being very uninformative. We want an interpretation of the framework of constitutive grounds which clarifies, in more positive terms, what might be involved when, say, attitudes help determine the good and bad.

A suggestion was explored according to which the constitutive grounds of final value are the things in virtue of which facts are afforded their normative status. In other words, the constitutive grounds of final value are the things that give facts the normative property of being a reason. It was then argued that this interpretation fits well with the general insights and observations that have often been made by proponents of subjectivism. Before the chapter ends, I wish once again to express my open-mindedness about some of the issues raised in this and the previous chapter. To reiterate, the concept of background conditions might be very wide in that it involves a large category of different phenomena. There is therefore no need to say which account of background conditions is the correct one, or which account would be most relevant when considering the kinds of substantive examples that we have been looking at. What we should say, at the very least, is that the hyperintensional analysis of background conditions captures one way in which final value can be affected by things other than value-makers and reasons. Similarly, although I believe that the interpretation of the framework of constitutive grounds presented here represents the best understanding at hand, I would not suggest that I have formulated the one true version of subjectivism. What I have highlighted is simply one of several ways in which one might be a proponent of this theory.

The biggest problem with the interpretation of the framework of constitutive grounds that I have been exploring throughout this chapter is its apparent failure in regards of neutrality. There is an intuition that the theories of subjectivism and objectivism should be neutral about many of the conceptual and substantive points that have been made throughout this chapter. It should not commit itself
to any contentious conceptual or substantive views. I have expressed sympathy with this objection and suggested that this chapter should perhaps be regarded as a challenge on part of proponents of the framework of constitutive grounds. The challenge is to formulate an informative interpretation of the framework which retains the desired conceptual and substantive neutrality, which upholds the apparent disagreement between subjectivism and objectivism, and which does not rely on further analogies and metaphors that are themselves in need of interpretation. I look forward to seeing whether the challenge can be met.
Chapter 6
On Evaluative Holism

We shall continue with investigations into the dependencies of final value by focusing on issues concerning its structure and behavior. Among the questions that we shall look very closely at are the following: How do the final value of complex wholes relate to the values of their respective parts, be they final or otherwise? Do parts retain the same final value irrespective of the context of their occurrence, i.e. regardless of what complex wholes they happen to be parts of? We shall see that these questions are not only of interest to the second-order domain as we here understand it, for it can also become relevant to our everyday reasoning about normative matters. If there is not a strong connection between the final value of complex wholes and the values of their respective parts, then it would seem foolish to assume that one such complex whole is finally better than another simply because it includes better parts.\footnote{By the same token, it would also seem foolish to assume that one complex whole is finally worse than another simply because it includes worse parts. Chisholm (1986, pp. 97–102) at one point offers up a striking example that is meant to illustrate this: Atheists and religious people alike have often been interested in the question how a benevolent God could allow bad things to happen. This is the problem of evil. If organicism is correct, then there is a possibility that the presence of bad things in a world can make the world as such better than it would otherwise have been. We shall return to some illustrative substantive examples later in the discussion, and especially toward the end of the chapter, where it will be used to highlight the features of the theories that will be our focus.} Likewise, if it turns out that parts may lose or acquire different final values merely because of the context of their occurrence, then we should be hesitant about drawing any general
conclusions about the final value of a part based solely on its apparent evaluative contribution to specific complex wholes.

In what follows, two different attempts at answering the two questions will be brought into focus: One attempt is represented by the theory of organicism, which states that the final value of complex wholes is not a function of the values of their respective parts. Among the many implications of this theory is that complex wholes can be good for their own sakes, even when they consist in nothing but bad parts. Another attempt at answering the questions is represented by the theory of conditionalism, on which parts may lose or acquire different final values depending on the context of their occurrence. This is compatible with parts being good for their own sakes in certain circumstances, even if they are bad for their own sakes in others. Besides establishing how reasonable and intelligible each of the views is, the main purpose of the following investigation is to gain clarity about what sort of dependencies there are between the values that we would intuitively place at distinct levels in the structure of the world. To this end, a lot of attention will be paid to how well the theories capture the link between final value and reasons for attitudes.

Proponents of organicism have occasionally defended the view that complex wholes can have final value without sharing any value with their respective parts (Moore, 1993/1903 p. 79). This means that there can be reasons to favor complex wholes for their own sakes and yet no reasons to favor their respective parts, just as there can be reasons to disfavor complex wholes for their own sakes and yet no reasons to disfavor any of their respective parts. Some philosophers think that this is strange. Dancy (2004) has suggested that because reasons to favor complex wholes are determined by the presence of their respective parts, we must have reasons to favor the parts as well. This means that the parts themselves must have value, lest the connection between value and reasons for attitudes be severed. The main problem with this objection is that it is unclear what motivates the insistence that because the reasons to direct attitudes toward complex wholes may be determined by the presence of their respective parts, we must also have reasons to favor the parts as well. On several understandings of the relevant notion of determination, there need not be this transmission.

After discussing the details of the objection just mentioned, we move on to a similar but more modest objection against organicism. If the theory is sound,

The theory is named and inspired by the principle of organic unities as it was defended by Moore (1993/1903, p. 79). However, we shall see that the two are distinct. It is possible to accept the principle of organic unities without thereby accepting organicism as it shall here be understood. We return to this issue later.
then it should be possible that there are complex wholes that do not share their value with any of their respective parts. Given the assumption that there is a connection between final value and reasons for attitudes, it follows that there could be reasons to favor complex wholes for their own sakes and yet reasons to disfavor some of their parts. Similarly, there could be reasons to disfavor complex wholes for their own sake and yet reasons to favor their respective parts. Brännmark (2001), Stratton-Lake (2002), Olson (2004), and Zimmerman (2014) suggest that aspects of this are problematic. The reason is not necessarily that there is a transmission of reasons from complex wholes to their respective parts, but rather that when we direct attitudes toward complex wholes, we cannot help but direct the attitude toward the parts that they are constituted by as well. In some cases, when the parts in question are sufficiently bad, doing so can be patently unfitting. It will be shown that on several understandings of how discerning fitting attitudes need to be, this possibility need not worry us either. To the extent that this is a legitimate worry, moreover, similar problems can be shown to apply for the theory of conditionalism as well (cf. Brown, 2007).

During the following discussions we shall see that conditionalism seems to follow straightforwardly from some of the assumptions made in this work. The theory also has the additional advantage of providing us with an attractive and structured picture of the world of values. In this regard, it appears to have significant theoretical advantages over the theory of organicism, which does not follow quite as straightforwardly from the understanding of final value on which we have so far built. However, by the end of the chapter, aspects of organicism will have been shown to at least fit well with that understanding and to offer up an intuitive explanation for certain cases of holistic behaviors in final value. Consequently, although there may be theoretical considerations recommending an exclusive choice between the two theories, I shall end by aligning myself with Hurka (1998) and McKeever and Ridge (2013) who are open to a hybrid view regarding the structure and behavior of final value. On the hybrid view, conditionalism and organicism could be regarded as complementary theories that can be combined to explain the unpredictable ways that values interact.

Remarks on the Right Kind of Parts and Values

The theories we will be looking at involve claims about the final value of complex wholes and its relation to the value of their respective parts. To start with we would do well with a rough idea of what these mereological notions
mean in the first place. In addition, we need an account of what specific parts and values should be the focus of the upcoming discussions. The need for such an account stems from the fact that it is unclear what parts and values are really relevant to the theories we will be looking at. Unfortunately, while I will be offering some general remarks on what complex wholes are and what it means for factors to be parts of such wholes, I shall not provide an account of what specific values should be the focus of the upcoming discussions. I will instead do what many philosophers tend to do in this context and simply stick my head in the sand. However, for the sake of my philosophical conscience, I wish first to acknowledge some of the problems that arise in the absence of such an account.

One easy start would be to say that a complex whole is a factor that decomposes into more than one part (cf. Simons 1987, pp. 26–28). Things become more difficult when we then look to the question what it means for something to be a part of such a whole in the first place. Indeed, this question turns out to be just as controversial as the issues that we have so far been primarily interested in. There are limits, then, to what we can possibly say about the notion of parthood without getting sidetracked. One observation we can make is that the varieties of parthood are manifold: Agency is part of personhood; ministries are parts of governments; antelopes are parts of herds; states of affairs are parts of possible worlds; puberty is part of growing up; speakers are parts of stereo systems; arms are parts of the body; hydrogen is part of water; milk is part of milkshakes; and so on. Philosophers have speculated that some of the examples in this list involve relaxed notions of parthood which should not be taken as literally as others. We will ignore such complications and approach the notion of parthood from a more evaluative perspective.

The notion of parthood that we are interested in involves a relation with the formal qualities of reflexivity, transitivity, and antisymmetry (Varzi, 2016). That the parthood relation is reflexive means that any object p is also a part of p; that parthood is transitive means that if p is a part of q and q is a part of r, then p is also a part of r; finally, that parthood is antisymmetric means that it entails neither symmetry nor asymmetry: If p is a part of q and p and q are identical, then there is a guarantee that q is a part of p as well. If p is a part of q and q and q and

---


216 For an overview of controversies surrounding the notion of parthood, see Varzi (2016).

217 This section is not about the conceptual analysis of parthood, but it takes a very general perspective in that it is about the notions of parthood that might be relevant to principles describing the structure and behavior of final value. Whether those notions might end up relaxed or not, from the perspective of discussions within other domains, does not concern me.
p are distinct from one another, then there is no guarantee that q is a part of p as well. If any of the examples mentioned in the previous paragraph involve notions of parthood that fail to display these formal qualities, then they are not relevant to the present discussions concerning the structure and behavior of final value, or so we shall assume. Now, there are reasons to think that these observations need to be amended with additional constraints that clarify what specific parts should be taken into consideration. The additional constraints are meant to help us avoid settling questions concerning the structure and behavior of final value in a seemingly arbitrary and premature manner.

Suppose we want to calculate the final value of the complex whole \( w \), consisting in the parts \( x_1, \ldots, x_n \). When we count these parts, we may wonder about the complex parts in \( w \) that are made up from the combination of two or more of the parts \( x_1, \ldots, x_n \). If the values of the more complex parts are to be counted in addition to the simpler parts \( x_1, \ldots, x_n \), then there is the risk that we will overestimate the final value of \( w \). One solution to this problem involves barring overlapping parts from being considered altogether. The parts \( x_1, \ldots, x_n \) overlap if, and only if, there is some factor \( z \) that is part of \( x_1 \) and \( x_2, \ldots, x_n \). Of course, once we have established that the parts \( x_1, \ldots, x_n \) do not overlap in this sense, we should also make sure that they do not leave room for gaps within \( w \). The parts \( x_1, \ldots, x_n \) would leave gaps within the complex whole \( w \) if, and only if, there is a part of \( w \) that is not \( x_1 \) or \( x_2, \ldots, x_n \). We otherwise face the risk of underestimating the final value of \( w \). The two additional constraints just laid out will require adjustment later, but for now we should take them for granted. Parts that satisfy these constraints will hereafter be referred to as ‘basic parts’.

Much more could be said about the notion of parthood, as well as the distinct kinds of complex wholes there are. However, I wish now to move on and address a much more neglected ambiguity here regarding the types of value that are relevant to the discussions at hand. Just as not all kinds of parts may do when we consider questions concerning the structure and behavior of final value, it seems safe to assume that not all kinds of value will do. This fact is often

---

218 There may well be several distinct decompositions of a complex whole that satisfies the two requirements just laid out and we may use any one of those decompositions to arrive at the final value of the complex whole. What we may not do is to add together the distinct decompositions and treat them all as contributing to the complex whole. Now, some philosophers have gone further than this and argued that there must be a privileged decomposition which should constitute the basis of evaluative calculations. This is important to note because this privileged decomposition is often described in terms of ‘basic parts’ as well. For more on how other philosophers use this term, see e.g. Harman (1967), Quinn (1974), Carlson (1996), Feldman (2000), and Zimmerman (2001).
overlooked by philosophers, but it seems to me just difficult as the one that has just been discussed. Organicism denies that the final value of complex wholes is necessarily determined by the value of their respective parts, but there is a question about precisely which types of value are to be considered. If we take an open-minded approach and count even the instrumental value accruing to parts, then the truth of claims like this will seem a bit trivial. After all, there is no end to the instrumental value that parts can have, and so it is not surprising that the value of complex wholes is not a function of them. However, if we take the close-minded approach and count only the final value that accrues to parts, then the truth of organicism may also turn out trivial, but for the opposite reasons.

To be clear, my worry is that the closed-minded approach would neglect important types of value, some of which could have special significance when it comes to questions concerning the structure and behavior of final value. We cannot immediately rule out at the start of our discussions that some of the neglected types of value might end up settling the extent to which complex wholes end up being good or bad for their own sakes. Bracketing them right at the outset would therefore be a big mistake. In the face of these kinds of issues, there seem to be at least three different options available: First, we could try to demarcate precisely which types of value are most likely to have the relevant connection to final value; second, we could try to look at the theories that shall be our focus and assess their plausibility against the backdrop of every single combination of different value types that we can think of; third, we could continue to do what philosophers have done so far in discussions concerning the structure and behavior of final value and sweep the present worry under the rug. At present, the third option seems very tempting to me, but I believe that there are some very minor demarcations that we can make before we proceed.

We shall be interested in the question how the final value of complex wholes is affected by the values of their respective parts. When considering this issue, we should have in mind the parts that are mutually exclusive, collectively exhaustive, and perhaps also basic. However, we should leave the question relatively open what types of value the parts need to have to merit our consideration. To start with, the only exception is made for instrumental value, which will be left out of the picture. We will also be interested in how the final

Irrespective of what kinds of parts or what kind of value we have in mind, perhaps the answers to questions concerning the structure and behavior of final value will turn out to be roughly the same. There might be reasons for optimism here. Eventually, there must come a time when philosophers end up being very lucky and things turns out to be just as simple as they have liked to assume.
value of parts is affected by the value of complex wholes. When considering this issue, we should also leave the question relatively open what types of value the complex wholes may have. However, we will once again start out by making an exception for instrumental value, which we should continue to ignore. Although the approach we will be taking ignores certain problems and complications when it comes to the structure and behavior of final value, some of them will be brought up and hinted at again when they seem very likely to affect the soundness of the arguments being considered.

The Substantive Failure of Evaluative Atomism

To get sense of the motivations behind the theories that shall be our main focus, we will begin by taking a closer look at evaluative atomism. This view makes two distinct claims: The first is that the final value of complex wholes is a function of the value of their respective parts. The second claim is that parts are necessarily such that they retain their final value irrespective of the context of their occurrence. Even though the two claims are meant to answer different questions concerning the structure and behavior of final value, they are here treated as constituting one compound view. This will make it easier to see that even if the claims are reasonable on their own their combination can lead to some absurd consequences. The problem is that evaluative atomism commits us to patterns of evaluation that should intuitively be left open or even rejected. Once the untenability of evaluative atomism has been demonstrated, it will be clarified what sorts of problems organicism and conditionalism must avoid for them to be considered plausible theories to begin with:

**Evaluative Atomism**: For any complex whole \( w \) and any decomposition of \( w \) into basic parts \( x_1, \ldots, x_n \), (i) it is necessarily the case that the final value of \( w \) is equal to some function involving the values of \( x_1, \ldots, x_n \), and (ii) it is necessarily the case that the values of \( x_1, \ldots, x_n \) are retained even outside the context of \( w \).\(^{220}\)

\(^{220}\) A proponent of classical hedonism might be tempted to accept evaluative atomism. The experience of pleasure always adds value to any complex whole in which it occurs and the extent to which the complex whole is valuable is determined by the number of experiences that adds value to it (as well as the strength and duration of these experiences, of course). This
This theory recommends that what should be considered to arrive at the final value of complex wholes are the values possessed by their respective parts when the parts are imagined across different contexts. Something more should perhaps be said about this issue. Moore (1993/1903, p. 187) famously suggest that to determine whether parts are good or bad for their own sakes, we need to imagine them as independent objects—divorced from any complex wholes they happen to be parts of.\textsuperscript{221} This sort of isolation test has since been shown to be implausible for several reasons. As Lemos (1994, p. 37) and Zimmerman (2001, p. 132) point out, entities like states of affairs cannot exist in absolute isolation. For there to be states of affairs in the first place, there must presumably be the objects, properties, and relations that are involved in and constitute the state of affairs in question. What is more, even the sparsest of possible worlds we can imagine will be populated by countless necessary objects, such as the formal truths of logic and mathematics. Because of the failures of the ontological version of the isolation test, Chisholm (1981) has gone on to suggest an intentional version of the same test, which is also defended in some form by Lemos (1994, p. 37) and Zimmerman (2001, p. 132). The intentional version of the isolation test puts emphasis on the act of contemplation, i.e. what we are to think of an object and what results from thinking of it in this way, normatively speaking. The idea is that an object has final value if, and only if, it would be fitting to direct an attitude toward the object for its own sake given that the object is contemplated as such. This intentional isolation test fits well with what has been said so far. One of the ways in which objects can be contemplated as such is by imagining what they would be like across different contexts. Although it is not without its flaws, this approach cuts through to the intrinsic properties of objects far more effectively than the ontological version of the isolation test ever could.\textsuperscript{222} Naturally, this means that the possibility of extrinsic final value cannot be accommodated, but this need not concern us

\begin{footnotesize}
\begin{enumerate}
\item A version of the ontological isolation test occurs in the discussions of Chisholm (1978). Roughly, he suggests that a state of affairs has final value if, and only if, i) there are possible worlds where that state of affairs represents all that is good or bad, ii) if the state of affairs is not neutral, then everything that represents all that is good or bad in the possible worlds either entails or is entailed by the state of affairs, and iii) the negation of the state of affairs also has final value. See also the objection raised by Bodanszky and Conee (1981).
\item There may be a worry that we will then bracket contingent value as well, i.e. the final value that accrues to the intrinsic but contingent properties of its bearers. However, this need not be the case. When we are to imagine parts across different contexts, we are to imagine that only the contexts are changing. Nothing in the intrinsic properties of the bearer can vary.
\end{enumerate}
\end{footnotesize}
now. All we are doing here is aiming to understand evaluative atomism and so complaining that it is incompatible with our assumptions concerning extrinsic final value would be stubborn and premature at this stage. For now, let us just note that evaluative atomism entails that all final value is intrinsic and return to the possibility of extrinsic final value later.

The understanding of evaluative atomism expressed by the definition naturally invites questions concerning the relevant function that is to be used to arrive at the final value of complex wholes. If the theory admits the possibility of there being more than one relevant function, then it would presumably need to be amended with a story about when certain functions are fitting rather than others and why. Needless to say, we cannot go into detail here about what such a story might look like. Let us therefore take the more traditional route and assume that evaluative atomism commits itself to the use of only one function, namely that of simple addition. To arrive at the final value of complex wholes, then, we are to add together the values of all their respective parts. Once this interpretation of evaluative atomism is in place and we look beyond any problems having to do with the possibility of extrinsic final value, then the theory will start to seem like a rather attractive one from a methodological standpoint. As Oddie (2001, pp. 314–318) points out, evaluative atomism has the theoretical advantage of providing us with a handy additive analysis. This additive analysis ascribes an order to the world of values with the aid of two distinct approaches:

\textit{Top-Down}: For any whole \( w_1 \) and any part \( x_1 \) in \( w_1 \), if there is a whole \( w_2 \) and a part \( x_2 \) in \( w_2 \) such that \( w_1 \) is better than \( w_2 \), then, \textit{ceteris paribus}, \( x_1 \) is better than \( x_2 \). For any whole \( w_1 \) and any part \( x_1 \) in \( w_1 \), if there is a whole \( w_2 \) and a part \( x_2 \) in \( w_2 \) such that \( w_1 \) is equally as good as \( w_2 \), then, \textit{ceteris paribus}, \( x_1 \) is equally as good as \( x_2 \).\footnote{These represent my understandings. For alternative reconsider Oddie (2001, pp. 314–318).}

\textit{Bottom-Up}: For any whole \( w_1 \) and any basic part \( x_1 \) in \( w_1 \), if there is a whole \( w_2 \) and a part \( x_2 \) in \( w_2 \) such that \( x_1 \) is better than \( x_2 \), then, \textit{ceteris paribus}, \( w_1 \) is better than \( w_2 \). For any whole \( w_1 \) and any part \( x_1 \) in \( w_1 \), if there is a whole \( w_2 \) and a basic part \( x_2 \) in \( w_2 \) such that \( x_1 \) is equally as good as \( x_2 \), then, \textit{ceteris paribus}, \( w_1 \) is equally as good as \( w_2 \).
Understood this way, the additive analysis of value puts quite strong constraints on the evaluative orderings that we can make. Once we have an evaluative ordering at any level of the structure of the world, many of the evaluative orderings that are left undone can be straightforwardly carried out downward or upward. Of course, this also means that if the additive analysis is sound, the way we reason and theorize about value is made significantly easier. The suspicion is awoken that to solve some of the more difficult problems in moral philosophy, all we need are numerical values with which to represent the weight of the relevant values to be added. For these reasons, it is perhaps not so surprising that aspects of the additive analysis, or, at any rate, something very much like it, have been used by philosophers to draw conclusions about controversial questions in moral philosophy. According to Oddie (2001, p. 317), one philosopher who does so is James Rachels. More specifically, Rachels appears to rely on something very much like the top-down approach in his influential discussions about the ethics of euthanasia, where he famously argues that there is no significant evaluative difference between killing a person and letting them die—assuming, of course, that the perpetrator is acting intentionally and knowingly. Just before he goes into the nitty-gritty of his argument, Rachels lays down the groundwork for his thinking by making the following analogy with what he considers to be scientific thinking:

The familiar procedure is to isolate the element of interest by studying cases in which everything else is held constant, while that one element is varied. Children are taught this idea at school by having them perform simple experiments. For example, does the colour of a combustible material affect whether it will burn? Children can see that it does not by trying—and succeeding—to burn bits of paper of different colours. Does the presence of air affect combustion? Most of us will remember placing a candle in a bell-jar and watching it go out after the oxygen is consumed, while a similar candle outside the jar continues to burn. By varying one element we see what difference it makes (1986, p. 112).

Rachels then applies this line of reasoning to two hypothetical cases: Smith and Jones each stand to gain a large inheritance if their respective cousins were to die. Both Smith and Jones therefore form the intention of murdering their respective cousins. The former does this by entering the bathroom where his cousin takes his daily baths and drowns him. He then makes sure that the incident looks like an accident and that none is the wiser. When Jones enters his
bathroom, however, he witnesses his cousin slip and hit his head against the side of a rail. All he must do then is stand back, cross his arms, and smile wickedly as his unconscious cousin drowns all on his own. According to Rachels, the only difference that we shall suppose between these two cases is the fact that one involves murder while the other involves accident and passivity. However, our assessments of the two cases tend to be the same and so Rachels concludes that the difference in question is unimportant: the act of killing and the act of letting someone die are in fact evaluatively equivalent. This could be understood as an application of something very much like the top-down approach.224

Regardless of whether we agree with his conclusions, the line of reasoning that Rachels makes use of brings the intuitive plausibility of the top-down approach to light. Applications of the bottom-up approach are arguably just as ubiquitous.225 Singer (1993, p. 222) at one stage proceeds from the result that there is no significant difference between killing and letting someone die in the case of euthanasia, to drawing conclusions about the contributions of such factors to cases involving the unequal distribution of material goods. Very roughly, if a person intentionally and knowingly lets another person die, when the aid of the latter would incur no significant cost to the former, this is just as bad as the former having intentionally and knowingly killed the latter. According to Singer, this shows us that letting someone die as a result of a lack of food, shelter, and clothes can be just as bad as actually killing them. Now, let us keep the additive analysis of value in mind as we move on to consider some objections to evaluative atomism. The purpose of those objections is to show that the additive analysis, when applied to a special type of cases, answers questions that should intuitively be left open and commits us to evaluative orderings that we may even disagree with.

224 Kagan is among those to criticize Rachels for precisely this reason, saying that he assumes that “the status of the act is the net balance or sum which is the result of adding up the separate positive and negative effects of the individual factors” (1988, p. 259). There are many other problems with the line of reasoning presented by Rachels, aside from its reliance on the additive analysis. For those who are interested in the discussions concerning the difference between killing and letting die that were partly ignited by the arguments of Rachels, see the overview in Woollard & Howard-Snyder (2016).

225 Anecdotally, it seems common within the context of everyday life as well. It happens quite often that a person is referred to as a hypocrite simply for treating the contribution of a factor differently when it occurs in different complex wholes. The tone in which such accusations are sometimes leveled seem to me to imply that being this kind of hypocrite is involved in a kind of incoherency.
Hints of Holistic Behaviors

With the claims and implications of evaluative atomism clarified in the previous section, we can now move on to look at it various problems. We shall consider whether the constraints that the additive analysis imposes on the evaluative orderings we can make is too strong. The worry is that we might be forced to accept one evaluative ordering based on another, when no such requirement seems necessary or intuitively plausible.\(^{226}\) This will not just show that evaluative atomism is problematic but will help us understand the alternatives that will be the primary focus in the discussion to follow. Finally, it will also grant us an opportunity to revisit the additional constraints on parts briefly outlined at the start of this chapter and show that they, too, can come into friction with our judgments. Let us start by considering the following complex wholes:

\[
\begin{align*}
C1: & \quad \text{Jones is happy and virtuous.} \\
C2: & \quad \text{Jones is unhappy and virtuous.} \\
C3: & \quad \text{Jones is happy and sinful.} \\
C4: & \quad \text{Jones is unhappy and sinful.}
\end{align*}
\]

Let us also assume that each complex whole in this list decompose into two of the following parts:

\[
\begin{align*}
C5: & \quad \text{Jones is happy} \\
C6: & \quad \text{Jones is virtuous} \\
C7: & \quad \text{Jones is unhappy} \\
C8: & \quad \text{Jones is sinful}
\end{align*}
\]

C1 is nothing over and above the parts C5 and C6; C2 is nothing over and above the parts C7 and C8; C3 is nothing over and above the parts C5 and C8; finally, C4 is nothing over and above the parts C7 and C8. Intuitively, it seems

\(^{226}\) The examples used in this section are borrowed from Lemos (1998; 2015).
to me that C5–C8 should not be barred from consideration when we want to calculate the value of their respective complex wholes. This intuition runs contrary to the constraint on overlapping parts mentioned earlier, since Jones in fact is a part of C5–C8. One strategy for handling this difficulty would be to add together the values of the parts C5–C8 and simply subtract the value of Jones whenever he reoccurs. Another strategy is to adjust the constraint on overlapping parts so that it only applies to those objects whose common parts are themselves the bearers of value. We can then simply stipulate a first-order value theory on which Jones lacks such value and get off relatively scot-free. For the sake of simplicity and concision we will employ the second strategy here.

Let us begin with the bottom-up approach and assume that while C5 and C6 are good, C7 and C8 are bad. We are also to suppose, as seems reasonable, that C5 and C6 are just as good as C7 and C8 are bad.\textsuperscript{227} The idea is to now use this evaluative ordering of the parts to arrive at the final value of the complex wholes C1-C4. The bottom-up approach yields an evaluative ordering on which C1 is finally better than C2. That C1 is better than C2 follows from the bottom-up approach because Jones being happy is better than Jones being unhappy and that is the only difference between C1 and C2. The bottom-up approach yields an evaluative ordering on which C3 is finally better than C4. That C3 is better than C4 follows from the bottom-up approach because Jones being happy is better than Jones being unhappy and that is the only difference between C3 and C4. However, my intuition is that that C4 must be finally better than C3,

\textsuperscript{227} Zimmerman (e.g. 2001, pp. 142–148) objects to many examples that are meant to illustrate holistic behaviors in final value, because of their involving cases of evaluative inadequacy. That the fact that Jones is happy is evaluatively inadequate means that it cannot be ascribed a definite value of either a positive or negative kind. The reason for its evaluative inadequacy is that the fact lacks a sufficient degree of specificity. For the fact that Jones is happy to have value, we must know more about the duration, and intensity of the experienced happiness. Perhaps we also must know what the happiness is ultimately about (Zimmerman, 2014, p. 13). However, once a fact becomes too specific, it will be evaluatively superfluous which also removes it from the prospective bearers of actual value (Zimmerman, 2001, p. 11). One problematic aspect of this objection is that it is difficult to cash out in a general way. The question is what decides when a fact escapes evaluative inadequacy—what level of specificity is required for it to be a bearer of value. Another problematic feature of the objection is that it appeals to intuitions that many, including myself, just do not share. The facts that are identified as evaluatively inadequate do seem to me to be possible bearers of value. After all, as Lemos (2003, pp. 589–590) notes, there can be reasons to favor and disfavor such facts for their own sakes (and there can be reasons to prefer one such fact over another for its own sake), even when though they do not exhibit the specificity that Zimmerman insists upon. This means that they can have value. At the risk of leaving readers somewhat dissatisfied, I will leave the objection aside for now. For more on evaluative inadequacy, see e.g. Lemos (1998, pp. 328–331; 2015, pp. 132), Hurka (1998, pp. 317–318), and Orsi (2015, pp. 90–93).
because the latter is a case involving undeserved happiness. This is reminiscent of the ideas expressed by Ross (2002, p. 138) on the topic of pleasure and sinfulness. He invites us to compare two possible worlds that are identical in terms of the values that they contain, but not in their distribution of these values. He states that if all else is equal, then that a possible world where all pleasure is had by the sinful and all pain is experienced by the virtuous is worse than a possible world where the reverse is true.

Even if we have reservations about the line of reasoning inspired by Ross, the result just reached should be enough to at least indicate that something is wrong with the bottom-up approach. We should be able to agree with the evaluative ordering of the parts with which we began and yet disagree with the evaluative ordering of the complex wholes we ended up with. Alternatively, if we feel very strongly committed to the additive analysis, we may instead backslide on the evaluative orderings of the parts and say that the part $C_5$ is bad when it occurs in complex wholes such as $C_3$. The idea behind this move is that the value of complex wholes can be understood in terms of the sum of the values of their respective parts, without the parts being subject to any sort of isolation test. To arrive at the values of complex wholes, then, we are to sum up the values that the parts happen to have within those complex wholes. Finally, the substantive failure of evaluative atomism can also be illustrated by way of the top-down approach, which move us from the evaluative orderings of complex wholes to the evaluative orderings of their respective parts.

Suppose we start out with an evaluative ordering of the complex wholes on which $C_1$ is better than $C_2$ and $C_4$ is better than $C_3$. The complex wholes $C_2$ and $C_3$ are arguably equal in value. The additive analysis entails that because $C_1$ is better than $C_2$ and the only difference between them is that $C_1$ involves happiness where $C_2$ involves unhappiness, happiness must better than unhappiness. So far so good. However, the additive analysis also entails that because $C_4$ is better than $C_3$ and the only difference between them is that $C_4$ involves unhappiness where $C_3$ involves happiness, unhappiness must be better than happiness. This means that we are seemingly faced with an outright contradiction in our evaluative ordering of the parts $C_5$–$C_8$. Now, the intuitive failure of the bottom-up and the top-down approach could be taken to indicate two things: Either there is something wrong with the additive analysis of value, or the values of $C_5$–$C_8$ shifts depending on their places in the complex wholes $C_1$–$C_4$. In other words, either the value of complex wholes is not an additive
function of the values of their respective parts, or the values of parts can come to depend on the context of their occurrence.\textsuperscript{228}

Oddie (2001, pp. 315–320) and Brown (2007, pp. 457–459) are among the philosophers to emphasize that although evaluative atomism entails the legitimacy of the additive analysis, the additive analysis does not necessarily commit us to evaluative atomism. More specifically, there are aspects of the additive analysis that can be preserved even if we deny that the values of parts are retained irrespective of the context of their occurrence. This is done by indexing the relevant parts so that they are not contemplated as such but are rather seen from the perspective of the specific complex wholes they happen to be parts of. It will no longer be true that to solve some of the more difficult problems in moral philosophy, we need little more than numbers and an additive function. In addition to these things, we also need an awareness of how the value of objects tends to be affected by the circumstance of their occurrence. This significantly reduces the effectiveness and usefulness of the additive analysis. What will result from the modification of the additive analysis under consideration are conditional approaches:

\textit{Conditional Top-Down:} For any whole $w_1$ and any part $x_1$ in $w_1$, if there is a whole $w_2$ and a part $x_2$ in $w_2$ such that $w_1$ is better than $w_2$, then, \textit{ceteris paribus}, $x_1$ is better in $w_1$ than $x_2$ is in $w_2$. For any whole $w_1$ and any part $x_1$ in $w_1$, if there is a whole $w_2$ and a part $x_2$ in $w_2$ such that $w_1$ is equally as good as $w_2$, then, \textit{ceteris paribus}, $x_1$ is equally as good in $w_1$ as $x_2$ is in $w_2$.

\textit{Conditional Bottom-Up:} For any whole $w_1$ and any part $x_1$ in $w_1$, if there is a whole $w_2$ and a basic part $x_2$ in $w_2$ such that $x_1$ is better in $w_1$ than $x_2$ is in $w_2$, then, \textit{ceteris paribus}, $w_1$ is better than $w_2$. For any whole $w_1$ and any part $x_1$ in $w_1$, if there is a whole $w_2$ and a part $x_2$ in $w_2$ such that $x_1$ is equally as good in $w_1$ as $x_2$ is in $w_2$, then, \textit{ceteris paribus}, $w_1$ is equally as good as $w_2$.

If evaluative atomism is true, then it is quite easy to see how the additive analysis has the potential of being very useful from a practical and theoretical standpoint. We can apply both the top-down approach and the bottom-up approach at any level of the structure of the world and get an accurate calculation of value,

\textsuperscript{228} See also Oddie (2001, pp. 322–323) who illustrates the same result with different examples.
assuming, of course, that we have an evaluative ordering on one level from which to start. However, if we accept the conditional modification of the additive analysis, this is not really the case anymore. We will no longer have a handy methodological tool for calculating value, although we are admittedly still given a nice and structured theoretical picture of the world of values. Let us linger on this point for a moment, for it can help us put the relative usefulness of the additive analysis into a proper perspective.

Suppose we use the top-down approach in the middle of the structure of the world, on a complex whole $w_1$ that is itself among the parts of a more complex whole $w_2$. The goal of applying the approach to $w_1$ is to figure out with certainty the values of whatever parts $w_1$ has. If we are proponents of evaluative atomism, no heed need be paid to the fact that $w_1$ is itself among the parts of $w_2$. Whatever result is reached from the application of the top-down approach on $w_1$ is guaranteed to be correct. However, if we accept the conditional versions of the additive analysis, we must make sure that we index the values of parts in the right way. We cannot look at $w_1$ in isolation and conclude anything certain about the values of its parts, because the value of $w_1$ will also have to be considered from the perspective of its context of occurrence. To determine the values of any parts within the structure of the world with certainty, we must start by establishing the value of a complex whole that is not itself a part and work our way all the way down to the point of interest. Similar issues crop up for the conditional version of the bottom-up approach.

Suppose we use the bottom-up approach in the middle of the structure of the world, on a complex whole $w_1$ that is itself also one of the parts of a more complex whole $w_2$. The goal of applying the approach is once more to figure out with certainty the value of $w_1$ based solely on the values of the parts that it has. Again, if we are proponents of evaluative atomism, no heed need to be paid to the fact that $w_1$ is itself one of the parts of $w_2$. The result we reach from the application of the top-down approach to the parts of $w_1$ is still guaranteed to be correct. By contrast, if we accept the conditional version of the additive analysis, we must make sure that we index the values of parts in the right way. We cannot simply look at the values of the parts within $w_1$ and conclude anything definite about the value of $w_1$, because the value of $w_1$ will also have to be considered from the perspective of its context of occurrence. Regardless of the contributions of its parts, $w_1$ can stubbornly defy expectations and change value depending on its place within the structure of the world. To avoid this kind of result, we must look at the parts of $w_1$ and index their values to $w_2$ and so on all the way up the structure of the world. Obviously, none of these issues constitute problems for
the theory of conditionalism, since they just illustrate the precise sense in which conditionalism is a kind of holism.

Let us step back and define evaluative holism as any view that affirms the falsehood of evaluative atomism. There are at least three different versions of holism available to us. First, we have organicism, which denies that it is necessarily the case that the final value of complex wholes is equal to some function of the values of its parts. Second, we have conditionalism, which denies that it is necessarily the case that the values of parts are retained even outside the context of their occurrence. Third, we have radical holism, which denies both claims made by evaluative atomism. This understanding takes evaluative holism to be constituted by compound theories as well and there is a good reason for this. It seems like a plausible suggestion that if the problems of evaluative atomism can be escaped by only rejecting one of its two claims, then, if all else is assumed to be equal, then this is preferable to rejecting both. There is therefore some theoretical pressure, however small, to avoid radical holism as a general story about the kind of phenomena we have just been looking at. For this reason, most of our attention will also be paid to understanding organicism and conditionalism, while, for much of the upcoming discussion, radical holism is treated as a last resort. We return to how we should best understand these theories in the upcoming section, where we will take a closer look at the mechanisms that seem to underlie them.

Now, just as it is reasonable to try to escape the problems of evaluative atomism by only rejecting one of its claims, it might seem reasonable to try to preserve as much as we can of the additive analysis. If this is right, then conditionalism starts out with some very strong points in its favor. For although organicism is also capable of avoiding the problems associated with evaluative atomism, it does so by giving up both on the top-down approach and the bottom-up approach. On organicism, no observations about the values of complex wholes can be used to infer anything certain about the values of their parts.

229 On this understanding philosophers that have been referred to in the literature as proponents of conditionalism, either by others or themselves, may in fact be radical holists. It is admittedly difficult to establish precisely which philosophers belongs to the category, since the distinction between conditionalism and radical holism is very seldom made. However, a small and incomplete selection of the many recent philosophers that have explicitly or implicitly accepted either conditionalism or radical holism would include Korsgaard (1983), Kagan (1998) Hurka (1998), Dancy (e.g. 1993, 2000, 2003, 2004), Rabinowicz & Ronnow-Rasmussen (2001, 2003, 2006), Olson (2004), and Orsi (2015). An equally small and incomplete selection of recent philosophers to reject either conditionalism or radical holism would include Chisholm (1986), Lemos (1998), Feldman (2000), Zimmerman, (2001), and Bradley (2002).
respective parts. The combination of conditionalism and the additive analysis, by contrast, paints a pretty picture of how values may interact at distinct levels of the structure of the world. This picture can help improve our understanding of value even if it does not make life much easier from a practical and theoretical standpoint. Whether these points are strong enough to make conditionalism more plausible than organicism remains to be seen. In an upcoming section, we shall see that the mechanisms that underlie aspects of both varieties of evaluative holism seem plausible and intuitive.\(^{230}\)

**On the Mechanisms of Evaluative Holism**

In this section, we should take a closer look at the varieties of evaluative holism and how they should be understood. To choose among the varieties of evaluative holism based on their methodological advantages alone would be premature. We need to get a better grip on the mechanisms that underlie them—the features of final value that would ensure that the theories in question could be accurate and sound. It might turn out that both theories are invited by features of final value that we have already recognized or committed ourselves to in the previous chapters. If this is so, then we will need stronger reasons to prefer one over the other than those that were hinted at in the previous section. We shall start out as we left off, by considering the mechanisms that underlie the theory of

\(^{230}\) Evaluative atomism may also be thought to fail for reasons having to do with the notion of *evaluative superiority*. The most famous defense of the notion can be found in the writings of Mill (1998/1863), wherein a distinction is made between higher and lower pleasures. Higher pleasures are achieved by way of intellectual endeavors, such as answering scientific questions or solving philosophical problems, while lower pleasures are achieved by paradigmatically hedonistic activities, such as eating or having sex. Mill thinks that no amount of lower pleasures would make a complex whole better than any amount of higher pleasures. In this sense, higher pleasures are not just better than lower pleasures, but they are in some sense insurmountably better. Several philosophers have since gone on to endorse the notion of evaluative superiority in some form or other (e.g. Edwards, 1979, pp. 69–72; Glover, 1977, p. 710; Griffin, 1986, pp. 85–86; Parfit, 1986, pp. 161–64; Crisp, 1988, p. 188; Lemos, 1993; Arrhenius, 2005; Arrhenius & Rabinowicz, 2005; and Rabinowicz, 2007). The worry here is that to make sense of evaluative superiority within the framework of evaluative atomism, the notion must be explained in terms of infinite value (cf. Ryberg, 2002). This would be problematic, but it is a result that an application of the fitting-attitudes analysis can help us avoid. We might say that x is a higher value in respect of y if, and only if, it is fitting to prefer any amount of x over any amount of y. To say this is not to suggest that x has an infinite value in the sense that it warrants attitudes of an infinite intensity or anything like that.
conditionalism. The question to be considered, then, is what the features of final value are that ensure that they can depend upon the context of its occurrence in the way suggested by the theory. Let us first remind ourselves of the following:

*Conditionalism*: For any complex whole $w$ and any decomposition of $w$ into basic parts $x_1, \ldots, x_n$, (i) it is necessarily the case that the final value of $w$ is equal to some function involving the values of $x_1, \ldots, x_n$, and (ii) it is not necessarily the case that the values of $x_1, \ldots, x_n$ are retained outside the context of $w$.

In the cases where our preferred evaluative orderings of parts turned out to be inconsistent with our preferred evaluative orderings of the complex wholes which included them, conditionalism entails that the final values of those parts changed in some way or other. The state of affairs that Jones is happy turns out to be a bad part of the complex whole that Jones is happy and sinful, while it turns out to be a good part of the complex whole that Jones is happy and virtuous. The question to be asked is how we should explain this behavior and what it is about final value that makes it so variable. One possible explanation that we should set aside here is that even in cases where final value depends upon the intrinsic properties of its bearer, it does not exclusively result from the essential features of its bearer.\(^{231}\) So, although the final values of parts are perhaps not directly sensitive to the circumstance of their occurrence, they may still change as long as the intrinsic properties of their bearers change.

This phenomenon can be illustrated by cases that we have already had occasion to consider in previous chapters. If a flower is exposed to too much sunlight, its intrinsic features can be irrevocably changed in the sense that it starts to wither. Once the flower has become brown and crumpled, it will have lost its beauty. What this means is that the context of a flower can have a causal influence on its intrinsic features, with the result that they are no longer capable of supporting its previous value.\(^{232}\) Unfortunately, although this story is certainly right as far as it goes, it is not quite strong enough to make conditionalism a successful response to the troubles that afflict evaluative atomism. We need to remember that proponents of conditionalism should want to claim that when

\(^{231}\) The possibility of final value resulting from contingent properties is discussed in e.g. Feldman (1998, pp. 352–353), Rabinowicz & Rønnow-Rasmussen (2000, p. 35), and Dancy (2004, p. 168). See also Dussault (2014).

\(^{232}\) What we referred to in chapter four as the ‘contributory background conditions’ undermine the value of the flower.
happiness is experienced by the sinful rather than the virtuous, the final value of happiness may have changed so that it is now bad for its own sake. The question, then, is what change has occurred in the intrinsic features of happiness when it stops being experienced by the virtuous and when it starts being experienced by the sinful. Without an informative answer conditionalism is left without the benefits that made it such a tempting theory to begin with. Furthermore, it seems intuitive that a proponent of evaluative atomism should be able to agree that final value does not just depend upon essential features.

At least two other explanations are available to proponents of conditionalism, both of which are found represented and defended in the previous chapters of this book. One such explanation relies upon the observation that final value can be sensitive to the extrinsic features of its bearers. In the current context, this should be taken to mean that the value-makers for a given final value do not have to be intrinsic features. We need not rehash all the examples that have already been given of this phenomenon. The point is simply that narrow normative explanations can include extrinsic content. Another explanation relies upon the observation that final value can be subject to background conditions, which may also be extrinsic relative to the relevant value-bearer. To reiterate, this means that there may be factors that explain, in a broad sense, why objects end up being good or bad for their own sakes, but without being intrinsic features of the objects themselves. Both the extrinsicality of value-makers and the sensitivity of intrinsic value-makers to background conditions afford a sturdy foundation to conditionalism. The theory of conditionalism is not just attractive because of its compatibility with the additive analysis, then, as it also fits very well with the assumptions and arguments made so far in this book.

Let us now consider whether the same thing can be said about the mechanisms of organicism. The question is whether there are features of final value that seem to make this theory accurate and sound as well. A brief glance at the history of philosophy reveals that, like many discussions in the second-order domain, contemporary debates concerning the structure and behavior of value were first ignited by the writings of Moore. He argued for the possibility that

---

233 The same goes for the first explanation which, though perhaps not strong enough to support conditionalism on its own, might appear reasonable enough to be adopted for its own merits.

234 This is why proponents of the idea that final value is intrinsic dismiss conditionalism. Moore at one point stated: "[I]t is impossible for what is strictly one and the same thing to possess that value at one time, or in one set of circumstances, and not possess it at another" (1922, p. 260).

235 As Chisholm (e.g. 1986, pp. 96–97) points out, however, the earliest contributions to discussions about the structure and behavior of value pre-date Moore. Even organicism—or at any rate, some limited version thereof—already occurs in the works of Aquinas and Brentano.
there can be mismatches between the value of complex wholes and the values of their respective parts. For example, there may be instances of complex wholes that are much better than the sum of the positive values of their parts, just like there can be complex wholes that are much worse than the sum of the negative values of their parts. He even thinks that there can be complex wholes with value that are nonetheless comprised of nothing but worthless parts. After elaborating on several possibilities along these lines, spelling out the many ways in which evaluative mismatches can occur, Moore concludes with his famous principle of organic unities: “The value of a whole must not be assumed to be the same as the sum of the values of its parts” (1993/1903, p. 79).

That brand of holism which is picked out by organicism is obviously akin to the principle of organic unities, but it is important to note that the two are not exactly equivalent. Perhaps organicism could be seen as affording us a particularly strong interpretation of the principle of organic unities. On this interpretation, it is not just that the value of complex wholes cannot be understood in terms of the sum of the values of their respective parts, but the value of complex wholes cannot be understood as any function of the values of their respective parts. Organicism thereby entails that the principle of organic unities is true, but the reverse is not the case. Another way of formulating the interpretation puts emphasis on the two approaches that were shown to characterize the additive analysis in the previous section. Organicism would then be understood in terms of a lack of constraints with respect to evaluative orderings. To make matters clearer, this is roughly how we will also understand the interpretation during the upcoming discussions as well:

**Organicism**: For any complex whole \( w \) and any decomposition of \( w \) into basic parts \( x_1, \ldots, x_n \), (i) it is not necessarily the case that the final value of \( w \) is equal to some function involving the values of \( x_1, \ldots, x_n \), and (ii) it is necessarily the case that the values of \( x_1, \ldots, x_n \) are retained outside the context of \( w \).

From a second-order perspective, any evaluative orderings we make at the level of parts will be compatible with whatever evaluative orderings we want to make

---

236 Of course, at several points in his discussions, Moore (e.g. 1993/1903, p. 79) also formulates this principle in terms of proportionality. He denies that the final value of complex wholes is necessarily proportional to the sum of the values of their respective parts. If the notion of proportionality is understood in a straightforward mathematical sense, this brings his principle of organic unities closer to organicism.
at the level of complex wholes. Of course, this is not to say that the final value of complex wholes is not dependent upon the presence of their respective parts themselves in any interesting sense. For example, even under the strong interpretation afforded by organicism, the principle of organic unities would be compatible with the assumption that the value of complex parts depends on the features of their respective parts. What this means is that there is a rather simple explanation on offer as to why mismatches in value can occur: Although the features of individual parts may sometimes fail to be value-making on their own, the features of distinct parts may serve as the resultance base for value once they are combined. Some philosophers might not think that this dependence is sufficient, and that organicism remains a pill that is too tough to swallow.

One suggestion is that even if the final value of complex wholes is not quite equal to the sum of the values of their respective parts, there must nonetheless be some function that can get us from one to the other. Perhaps the values of complex wholes are proportional to the values of their respective parts in a more sophisticated sense than we have so far considered. There are several reasons for why proponents of the principle of organic unities should adopt organicism, rather than hold out for a less obvious function in this way. One such reason is that no talk of functions is very likely to provide us with an explanation for cases where complex wholes that are good for their own sake consist in nothing but bad parts. It should also be mentioned here that organicism could be interpreted by reference to the final value of the mere arrangement of parts (e.g. Dancy, 2003, pp. 635–636; Zimmerman, 2001, pp.136–137). Such an interpretation is to be avoided when possible, because it involves more cumbersome explanations if it is to accommodate the things that proponents of organicism want to accept.

Suppose some good parts may come together and constitute a complex whole that is bad for its own sake. In such a case, the arrangement interpretation suggests that it must be because the arrangement of the good parts has a negative value that outweighs the positive final value of their sum. Conversely, when bad parts come together and constitute finally good complex wholes, it must be because the arrangement of the parts happens to be better than the negative final value of their sum. Whenever worthless parts come together to constitute a complex whole with final value, this is only because the arrangement of those parts is either good or bad. What is even stranger is that as soon as parts with value ever come together to form a worthless complex whole, the arrangement of the parts in question must be either good or bad to the precise degree that it in effect neutralizes the sum of the final values of the parts. The arrangement interpretation may help us avoid the problems of evaluative atomism, but
without the interpretation in question organicism seems to me to grant us more possibilities when it comes to the evaluative orderings that we can make. 237

In the current context, the most problematic aspect of organicism seems to be its second claim, according to which the final value of parts is retained irrespective of the context of their occurrence. This does not fit well with the assumption that it is possible for an object to have final value in virtue of, or as a result of, its extrinsic features. Nor does it fit well with the view that final value is sensitive to background conditions, a possibility which was defended at length in previous chapters. However, although organicism may not work as a general claim about the structure and behavior of final value, it might hold true for those final values that do not depend upon extrinsic features in any of the senses just mentioned. In other words, it might be that the second claim of organicism makes perfect sense for all final value that also happens to be intrinsic value. Toward the end of the chapter, I will therefore suggest a hybrid view according to which conditionalism and organicism could be regarded as complementary theories regarding the structure and behavior of distinct kinds of final value. Until then, I will be focusing mostly on the first claim made by organicism and the many criticisms that it has provoked from adherents of its rival views. Although this claim also turns out to be rather controversial, I will try to show that it is intuitive both from a formal and a substantive point of view.

The question might be asked is whether the attitudes we have reasons to direct toward complex wholes must match both the polarity and intensity associated with the sum of the attitudes that there are reasons to direct toward their respective parts. Let us imagine that we have before us a complex whole \( w_1 \) comprised of the basic parts \( x_1, x_2, \) and \( x_3 \). Let us also assume that the weight of value, and that the corresponding intensity of attitudes, can be represented numerically. In the context of \( w_1 \) each of the parts \( x_1, x_2, \) and \( x_3 \) has a positive value calling for attitudes with an intensity of, say, +10. This means that if the attitudes that all the parts call for have a combined intensity of +30. It seems to me like an open question whether the attitude we have reasons to direct toward \( w_1 \) would necessarily have to match this combined intensity of +30. On my view, that there should be some feature of value that guarantees such a match would be the truly surprising result. There simply is no ground, at least from a

---

237 For similar reasons, we also leave aside any interpretation of organicism whereby the final value of complex wholes is seen as a function of the values of their respective parts taken together with the values that the complex wholes themselves have as the whole. Moore (1993/1903) himself seems to have been a proponent of this view. For a discussion on this and related issues, see Carlson (2001, 2015).
second-order perspective, for saying that the attitudes we have reasons to direct toward complex wholes must match both the polarity and the intensity associated with the sum of the attitudes that there are reasons to direct toward their respective parts. Nevertheless, some philosophers have argued that if we keep the connection between value and reasons for attitudes closely in mind, then organicism turns out to be problematic. I turn now to their objections.

**A Transmission Principle for Parts and Wholes**

We should now have a sufficiently tight grip on the relevant theories that we can take a closer look at their respective benefits and disadvantages. Conditionalism seems to have a significant theoretical advantage to start with, not only because it affords a structured view of the world of values, but also because it fits well with the assumptions made in this work concerning. To reiterate, the assumptions in question are about the extrinsicality of value-makers and the sensitivity of value-makers to background conditions. However, we have also seen that there are mechanisms that ensure that parts of organicism could also be accurate and sound. In this and the upcoming section, we shall consider whether there is any argumentative leverage to be found against that theory in the connection between value and reasons. The first objection we shall consider relies upon the idea that complex wholes necessarily transfer reasons to their essential parts. What this means is that if there are reasons to favor a complex whole for its own sake, then there are reasons to favor its essential parts as well. The idea that there can be complex wholes with final value that consist of worthless parts is therefore false. The style of objection first occurs in the writings of Dancy. It would be helpful to take our point of departure from the formulation that he provides, although we should be careful to treat the objection in a general way and not get too stuck in the general interpretative issues that may crop up here:

[T]o say that a part with no value can contribute value that it has not got commits one to saying, it seems, that though there is no reason to preserve the part as a part, there is a reason to protect the whole, and that reason derives from the presence of the part. Now this does sound incoherent. Surely we do have reason to protect the part here, if it is contributing value. So its presence is of value, it would seem, on pain of breaching the link between values and reasons for attitudes (2003, pp. 630–631).
If we take the point to be about whether the final value of complex wholes confer final value on their respective parts, then we can dismiss this argument rather easily. For it is obvious that even if the final value of complex wholes confers some type of value on their respective parts, it does not follow that the value accruing to the parts is also final. However, this is not how we should understand the dialectic here. The question is whether the final value of complex wholes must confer any type of value of their respective parts—whether the reasons we have to favor complex wholes for their own sakes has any implications for whether we have reasons to favor the parts. This question is of a general interest and the argument by Dancy provides us with a good opportunity to look closer at its possible answers. The view that I will ultimately be defending is that even if complex wholes have final value, this does not necessarily have any implications for the value of their respective parts.

In the quotation Dancy expresses doubt about the idea that parts can contribute value to complex wholes that the parts themselves lack. He later adds to it by stating: “If a part contributes value, it must have at least as much value as a part as it contributes” (2003, p. 635). One of the things that make this argument a bit difficult to follow is the vague notion of a part contributing value. It needs to be clarified exactly how this notion should be understood. Parts contribute to the value of complex wholes in so far as they help determine the presence of the value. On my understanding, the reasons to favor a complex whole can be determined by the presence of parts in multifarious ways, among which are the following: First, it may be that the reasons to favor the complex whole are facts about the collective features of its parts. This means that if the parts in question are absent, then the facts that constitute the relevant reasons might not be instantiated. Second, it may be that the collective features of the

---

Dancy acknowledges that there are many different views regarding the connection between values and reasons for attitudes. He claims that what is required for his argument to go through is merely that there is a necessary biconditional between propositions about what is valuable and propositions about what we have reasons to favor. Of course, this means that Dancy overstates things quite a bit when he claims that his argument could only be undermined by “a position that maintained that there need be no connection between values and reasons” (2003, pp. 630–631). As McKeever and Ridge (2013, p. 275) note, proponents of organicism may claim that it is a matter of a much weaker entailment thesis such that wherever there is value there are also reasons for attitudes, but not vice versa. Later in their discussion, McKeever and Ridge (2013, pp. 273–276) also develop a response on which many instances of holistic behavior in value are cases involving attributive value, which they assume are not to be understood in terms of the presence of reasons. Significant effort was expended in chapter 2 to show that this assumption is incorrect. Be that as it may, we shall hereafter ignore any possible rejoinders that try to get around the objection by Dancy by denying that there is a strong connection between value and reasons to start with.
relevant parts explain why certain facts are reasons to favor the complex whole. This means that if the parts in question are absent, then the relevant facts might not constitute reasons. Third, it may be that the relevant parts are essential with respect to the complex whole. This means that if the parts in question are absent, then the complex whole that we would otherwise have reasons to favor would not exist. Let us go through each of these understandings in turn.

Suppose the reasons to favor a given complex whole are determined by the presence of its parts in the sense that the reasons in question consist in facts about the collective features of some of those parts. This means that the collective features of the parts serve the role of makers with respect to the value of the complex whole in question. Dancy appears to have something similar in mind when he later tries to formulate his objection without relying on the vague notion of a part contributing value. He says that “if a part is an element in the ground for the value of the whole, rather than playing some other role, perhaps as an enabling condition, the value of the whole must be explained partly by appeal to the value of the part, not merely by appeal to its presence” (2003, p. 634). Although this formulation is clearer than the one he gives in the original quotation, it seems to rely on two problematic ideas: First, just because the collective features of parts serve as makers for the value of a complex whole, we cannot draw the conclusion that the parts themselves serve the role of makers. Second, just because the collective features of parts serve as makers for the value of a complex whole, we cannot draw the conclusion that the collective features of parts are themselves bearers of value.

As Zimmerman (2001, p. 19) and Lemos (2005, p. 183) note, we should not generally assume that just because an object has final value in virtue of, or as a result of, certain features, the features themselves must also be good or bad. Value-makers need not be value-bearers. The same idea seems to hold true for parts in the sense that we should not generally assume that just because an object has final value as a result of certain features, the presence of the objects bearing those features must also be good or bad. The bearers of value-makers need not be value-bearers. Let us imagine that we ask a hedonist to give a narrow normative explanation why pleasure is good for its own sake. Her answer is that it has final value in virtue of, or as a result of, the particular subjective quality that is associated with the experience of pleasure.239 We insist that the particular subjective quality that the experience of pleasure has must therefore also be good for its own sake. We therefore need to understand why the peculiar subjective

239 What makes pleasure good for its own sake is “what it is like” to feel pleasure, to borrow a famous phrase from Nagel (1974).
quality that the experience of pleasure has is good for its own sake. At this stage, I think that the hedonist would be right to put her foot down and insist that the buck stops at the felt quality of pleasure. Formally, nothing more is required of her. Moore appears to make the same point when he writes:

It has just been said that what has intrinsic value is the whole, and that this includes the existence of the part; and from this it would seem a natural inference that the existence of the part has intrinsic value. But the inference would be as false as if we were to conclude that, because the number of two stones was two, each of the stones was also two (1993/1903, p. 81).

That value-makers need not be value-bearers is a notion that can perhaps be given further support by the more general observation that complex wholes often have features that are not shared by any of their respective parts, even though they are dependent upon the presence of these parts. This is true of the relation between the features of groups and the features of their respective individual members. There appears to be no inconsistency in the claim that even though the higher-order properties of a group of individuals depend upon the presence and features of their individual members, none of those members share any of those higher-order properties as individuals. The question, then, is why it would betray any confusion to make the parallel claim that even though there are reasons to favor a group of people which depend on the features of its individual members, we have no reasons to favor any one such member. The claim that it would appears particularly strange given the fact that at the most basic level most of the things that we favor for their own sakes presumably consist of things that we do not. This should not be taken to show that organismic is true, lest the theory seem a bit trivial, but, at the very least, it illustrates that the phenomenon Dancy finds peculiar is quite ubiquitous.

Let us move on to the second understanding of how the value of a complex whole can be determined by the presence of its parts. Suppose that the collective features of the relevant parts explain why certain facts constitute reasons to favor the complex whole. This means that if the parts in question are absent, then the relevant facts might not constitute reasons. According to the understanding developed in previous chapters, this means that the collective features of the relevant parts might be seen as background conditions for the relevant value. Regarding this possibility, the same arguments that have just been given seem to me to apply. We should not generally assume that just because an object has final value that depends upon the presence of certain background features, that
those features must themselves be good or bad. Background conditions for value need not be value-bearers. Finally, we should not generally assume that because an object has final value that depends upon the presence of certain background features, that the presence of the objects bearing those features must also be good or bad. The bearers of background conditions need not be value-bearers. Let us linger on this point for a moment longer.

Imagine next that we ask someone to give a broad normative explanation why a joke made at the expense of a colleague is humorous. We notice that in the broad normative explanation that we are given, the collective features of parts are mentioned. Those collective features are meant to explain why certain facts end up being reasons to be amused by the complex whole in question. It may therefore be tempting to insist that the collective features themselves must be good or bad, but it seems as if this can very easily be resisted. For what goes for the makers of value also goes for the makers of reasons. Interestingly, McKeever and Ridge (2013, pp. 279–280) have recently put forward an argument which could be taken to show that even Dancy needs to resist this interpretation of what he says. To be specific, their argument is meant to show that the objection that Dancy develops could be taken to undermine his own views on the concept of background conditions. They focus on is the question why the value of complex whole transfers value on the parts that are the resultant base for the value, but not on the parts that play the role of background conditions. They formulate the argument as follows:

In the case of parts, however, Dancy claims that parts do contribute value (and clearly thinks a defender of organic unities must agree). This claim, however, seems to turn on the fact that the parts are a means to value. Since one can equally well say that enablers are a means to value, his argument here commits him to saying that whenever something enables something else to contribute value, that enabler itself ipso facto contributes value. That seems analogous to saying that whenever some consideration enablers another consideration to favour some response (that is, to be a reason), that enabler itself favours that response (that is, is a reason in favour of it) (2013, p. 279).

Let us move on to the third understanding of how the value of a complex whole can be determined by the presence of its parts. Suppose we have a case before us such that if the parts of a complex whole with final value were to disappear, then so would the complex whole that they constitute. The question is whether this
means that there must also be reasons to favor the parts themselves. Although he
does not explicitly invoke it, perhaps the style of objection that Dancy develops
could be seen as relying on a transmission principle on which complex wholes
with final value transfer reasons to their essential parts. The idea is that when
complex wholes have final value, this does not entail that there are reasons to
hold attitudes with respect to all their parts, but only those whose presence is
required for the existence of their respective complex wholes. This transmission
principle is akin to that more well known transmission principle in second-order
theory which states that reasons transfer from ends to their necessary means.²⁴⁰

For the sake of illustration, suppose we have reasons to go to Stockholm. If it
turns out that buying a train ticket to Stockholm is the only means by which we
may eventually go to there, then it seems reasonable to insist that we have
reasons to buy a train ticket. In other words, if someone acknowledged that
there are reasons to go to Stockholm and that buying a train ticket is the only
means of doing so, then it would be strange if they went on and denied that they
had any reasons to buy a train ticket.²⁴¹ The relation between ends and necessary
means is certainly not identical to the relation between complex wholes and
essential parts, but it might be thought that they are sufficiently alike that
reasons would behave similarly for both relations. Side by side, the two
transmission principles would look something like the following:

²⁴⁰ For a discussion of this transmission principle in the case of reasons, oughts, or rationality, see
e.g. Kolodny (forthcoming), Raz (2005), Setiya (2007), Schroeder (2007, 2009), Bratman
²⁴¹ There are different versions of the transmission principle that vary in strength (Schroeder,
2009, p. 245). When applied to ends and means, the weak version says that if there are reasons
to \( p \) and \( q \) is a necessary means for us to \( p \), then we have reasons to \( q \). The strong version
would say that if there are reasons to \( p \) and \( q \) is a necessary means for us to \( p \), then we have
reasons to \( q \) that are just as strong as our reasons to \( p \). These versions of the principles can be
made to apply to the present context as well. In the case of complex wholes and parts, the weak
version would then say that if there are reasons to favor \( p \) and \( q \) is an essential part of \( p \), then
there are reasons to favor \( q \) as well. The strong version would say that if there are reasons to
favor \( p \) and \( q \) is an essential part of \( p \), then there are reasons to favor \( q \) that are just as strong as
the reasons to favor \( p \). Now, at first sight it might be thought that these versions of the
transmission principle speak in favor of the examples given by Moore. When applied to
complex wholes and parts, the weak version seems to allow for the sum of the constitutive
values of essential parts to be weaker, stronger, or equally strong as the value accruing to their
respective complex wholes. Similarly, the strong version may seem to allow for the sum of the
constitutive values of the essential parts of complex wholes is always stronger than the value
accruing to the complex wholes themselves, if the complex wholes have more than one
essential part. However, we should recall that the strength of values is to be analyzed in terms
of the intensity of the attitudes we have reasons to have and not the strength of the reasons.
**IT:** If we have reasons to pursue an end, then we have reasons to pursue the necessary means toward that end. In other words, if we have reasons to \( p \) and \( q \) is a necessary means toward \( p \), then, necessarily, we also have reasons to \( q \).

**MT:** If we have reasons to favor a complex whole, then we have reasons to favor the essential parts of that complex whole. In other words, if we have reasons to favor \( w \) and \( p \) is an essential part of \( w \), then, necessarily, we also have reasons to favor \( p \).

If IT is correct and ends transfer reasons to their necessary means, then perhaps it is only reasonable to suppose that MT is also correct and that complex wholes transfer reasons to their essential parts. Let us suppose for now that both transmission principles are true. If there are reasons to favor a complex whole for its own sake, then we must have reasons to favor its parts as well. This invites the question what kind of value would be entailed by the presence of the transferred reasons. The analogy with ends and necessary means could be taken to indicate that the reasons support an instrumental value. Complex wholes with final value would confer instrumental value upon their essential parts. A proponent of organicism may perhaps acknowledge this without much cost to their views, since instrumental value is one of the things that objects come by rather easily. However, this is a point at which the analogy with means and their necessary ends breaks down. The relation between complex wholes and their essential parts is a constitutive relation and so the final value of complex wholes transfers constitutive value to their essential parts. The essential parts are to be favored for the sake of their constitutive role and not for the sake of their effects.\(^{242}\)

The strong interpretation afforded by organicism claims that there are no necessary connections between the evaluative orderings of parts and the evaluative orderings of the complex wholes that include those parts. If what has

\(^{242}\) It is worth pointing out that like instrumental value, constitutive value seems to be extrinsic in the sense that it does not just result from the intrinsic features of its bearers. Constitutive value is an extrinsic value of a different kind, however, since the causal powers of essential parts are not necessarily part of the normative explanation as to why they have value. Now, we have seen in previous chapters that the concept of final value is one which admits to different forms. Perhaps it is possible, then, for constitutive value to also be of the final kind. However, this is something that proponents of organicism may want to reject, lest they become radical holists. Their view is that holistic behavior should not primarily be understood in terms of conditionalism. There is a possibility, then, that proponents of organicism can shrug off the objection and say that their view is silent on instrumental value and constitutive value.
been said is true, then this turns out to be false with respect to what we consider to be essential parts, which must be ascribed constitutive value. One response would be to reject the strong interpretation afforded by organicism and understand the principle of organic unities literally, as it was originally formulated by Moore. The claim that the final value of complex wholes cannot be understood in terms of the sum of the values of their respective parts is not shown to be false by pointing out that complex wholes confer constitutive values on their essential parts. This means that although the examples which Moore finds so appealing might need some qualifications and amendments, his general views about the structure and behavior of value seem to be largely preserved in spirit. This is certainly right as far as it goes, but my feeling is that this option would be a bit premature at this stage. In what follows, we shall consider two reasons why we might want to reject mereological transmission. Then we shall also consider a more general reason why we might want to reject the principles expressed by both IT and MT.

Suppose again that we have reasons to go to Stockholm which result in our having reasons to buy a train ticket. The reasons are meant to transfer from our going to Stockholm to our buying a train ticket, because unless we do the latter we cannot do the former. If we eventually do go to Stockholm, then it follows that we must have bought a train ticket. When this is emphasized, the relations under consideration will start look very different from one another. Reconsider the case where we have reasons to favor a complex whole that consists in essential parts. The reasons are meant to transfer from our favoring the complex whole to our favoring the essential parts, but here it seems obvious that we can do the former without doing the latter. Suppose the complex whole is a possible world that is identical to this one, save for one positive difference: The inhabitants of that possible world have just found a cure for cancer. Otherwise everything is identical to the actual world. If the relevant responses are understood in terms of attitudes, like the attitude of wishing, then it seems clear that we need not favor every part of that possible world to favor the possible world as a complex whole. We shall return to this point in the upcoming

---

244 It should be noted that Dancy might be okay with this. He says that parts cannot contribute value that they lack but admits to the possibility that some parts can fail to contribute value that they have: “So a feature can contribute less value than it has, but not more” (Dancy, 2003, pp. 667). In this sense, he acknowledges some of the organic behavior of value.
sections, where I shall distinguish between two ways in which objects can figure in the intentional contents of attitudes.

Relatedly, suppose we accept the view of mereological essentialism and so believe that all parts are essential to their complex wholes (e.g. Chisholm, 1973, pp. 581–582). The view implies that as soon as one part is replaced with another in a complex whole, then what we have before us is an entirely new complex whole. It is not difficult to see what it would mean if mereological essentialism was combined with the transmission principle: We would then have it that complex wholes with final value confer constitutive values on literally all their parts. Proponents of the transmission principle could obviously get around this result by denying mereological essentialism, but this is not plausible as a general strategy. The problem is that the view is trivially true for certain kinds of complex wholes: Facts, states of affairs, propositions, and possible worlds, are such that their parts cannot be replaced without resulting in new facts, states of affairs, propositions, and possible worlds. It seems to me violently counterintuitive that the ascription of final value to something like a possible world should necessarily commit us to the ascription of value on literally all its constitutive parts just because they are all necessary. To put things stubbornly, there simply is no such commitment.

We shall return to examples of this kind soon enough, but before that let us also consider a more radical response on which both mereological transmission and instrumental transmission ought to be rejected. Broome (2013, p. 126) argues that whether oughts transfer from ends to means can come to depend upon our intentions. This line of thought is applicable to the case of reasons as well. If we assume that we have no intentions whatever of going to Stockholm, and that we know that we will not end up going, then it seems bullheaded to insist that we still have reasons to buy a train ticket. The suggestion that someone who disagreed with this conclusion would have to be incompetent regarding the concept of a reason seems to be to become rather tenuous. If this is right, then perhaps the same response can be applied to the case of complex wholes and their essential parts. The former would then transfer reasons to the latter only in so far as we have the right intentions. Suppose we have reasons to favor some complex wholes. If we have no real intention of ever

245 Kiesewetter (2015) has worries that this line of reasoning would rely upon the theory of actualism, which states, roughly, that our reasons to commit an action are partly determined by the actual outcomes of that action. The theory was first laid out in Jackson and Parfit (1986). I do not think that the line of reasoning would rely upon the theory of actualism, but I will not go into depth about this here.
complying with our reasons to favor the complex wholes, perhaps because we have identified strong reasons that speak against doing so, then the claim that we must have reasons to protect or preserve their respective essential parts begins to seem rather questionable. In my view, this suggestion is no stranger than the notion that facts about reasons coupled with facts about our intentions can figure together in the explanation why facts constitute reasons for attitudes to begin with.

**Unfitting Ambivalences and Normative Asymmetries**

The argument presented by Dancy, on which the final value of complex wholes transmits value to their respective parts, seems to me to fail. He does not manage to show that organicism or the principle of organic unities comes into conflict with the close connection between value and reasons. Perhaps there are other interpretations of what it means for complex wholes to be determined by the presence of their parts that are relevant to the previous discussion. Some of those alternative interpretations may be used to strengthen Dancy’s original objection to organicism and the principle of organic unities. I therefore settle for the modest conclusion that none of the interpretations considered so far seem very helpful in this regard. Nevertheless, it may be argued that the objection formulated by Dancy can be modified to establish a weaker point than he originally intended. The point is that organicism and the principle of organic unities commit us to the possibility of some very troubling normative asymmetries and unfitting ambivalences. An objection seemingly along these lines is put forward by Brännmark (2001), Stratton-Lake (2002), Olson (2004), and Zimmerman (2014). The objection they have in mind proceeds from observations about radical mismatches where complex wholes are good, even though, according to organicism, the complex wholes consist in objects that are very bad. Let us start by considering the following:

\[C9:\] John is pleased by the fact that Mary is in pain.

\[C10:\] John is pained by the fact that Mary is in pain.\textsuperscript{246}

\textsuperscript{246} The examples are borrowed from Olson (2004), but it should be noted that it is problematic. Intuitively, C9 does not decompose into the parts ‘John is pleased’ and ‘Mary is in pain’, for the pleasure that John experiences is about Mary’s pain. The same observation holds for C10. We will be generous and leave this problem aside for now. Even if there is a very similar
For the sake of argument, we are to suppose that pleasure is good while pain is bad. Pleasure is part of C9 whereas pain is part of both C9 and C10. Even though C10 includes one part that is bad and no part that is good, the whole is a case of compassionate pain which is good. Olson (2004, pp. 37–43) argues that we cannot have reasons to favor C10, because it includes bad parts. The part he worries about is that Mary is in pain, which there are arguably reasons to disfavor. Olson thinks that it is much more intuitive to say that the holistic behavior apparent in these types of cases is to be explained by reference to the variability of value. It is important that we get this objection right. Olson could admit that even though there are reasons to favor complex wholes, there are no reasons to favor their respective parts. On his view, the problem is rather that there can be parts that are so bad that directing a positive attitude toward their corresponding complex wholes will seem to be unfitting. The fact that Mary is in pain is meant to be a case in point. Olson states that directing a positive attitude toward a complex whole that includes the part in question is too “indiscriminate” (2004, p. 39).

The account that Olson prefers would not try to explain the holistic behavior apparent in this case by looking at the value of C10 as such. The fact that John is pained has positive value within the context of C10 and therefore calls for positive attitudes. The fact that Mary is in pain has negative value and therefore calls for negative attitudes. Exactly how Olson wants to deal with the value of C10 as such is admittedly a bit unclear, but the main point is that the relevant value that results from the combination of John being in pain and Mary being in pain is in the former. The fact that John is in pain becomes very good. I

---

247 Interestingly, Olson (2004, pp. 37–42) appears to understand the relation between values and reasons for attitudes as a mere entailment, but he prefers to formulate it in terms of a connection between values and appropriate attitudinal responses. This minor difference will not matter in the present context. Though we have chosen to formulate our working assumption in terms of value and reasons, the arguments we will be considering apply just as well to accounts that focus on other concepts in the deontic family.

248 It is interesting to note, here, that although Zimmerman (2014, pp. 12, 20–21) agrees with this objection when it comes to C9, he is far less confident about the objection when it comes to C10. Here he thinks that the relevant ambivalence would be fitting.

249 In fact, Olson appears to be a proponent of radical holism. Regarding the kinds of complex wholes that were listed above, he suggests at one point (2004, p. 42, n. 31) that the relevant values are heterogeneously distributed among their respective parts in a way that cannot be characterized along the lines of any additive analysis. His adoption of radical holism seems motivated by its economy. Olson argues, in a last section of the paper, that his view, as opposed to organicism, does not commit itself to ascribing contributory value to parts in
think that this is a stronger argument than the one formulated by Dancy. However, it seems to me intuitive to respond by suggesting that when an attitude is directed toward complex wholes, then the parts of the complex wholes can be represented in the attitude in a less direct sense than the complex wholes themselves. I think that this provides us some leverage against the objection under consideration. A similar point is made by Lemos (1994, p. 10) when he distinguishes between loving objects *simpliciter* and loving objects *per accidens*. He notes that even if one loves certain objects *simpliciter*, it does not follow that one loves any of their respective parts in the same way, for these parts might just be loved *per accidens*. I wish to linger on this point, since it could turn out to be central to value and its connection to reasons for attitudes.

The distinction between loving objects *simpliciter* and loving them *per accidens* is meant to capture the fact that factors can figure at distinct levels in the intentional contents of attitudes. Parts can be represented in attitudes in a rather peripheral way that mainly serves to characterize or identify their respective complex wholes. In such cases, it seems misleading to suggest that the parts themselves are what the attitude is directed toward. There is, at the very least, one obvious sense in which the attitude is not aimed at the parts in question. Analogies practically write themselves at this point: When an archer notches an arrow and aims it at a bullseye, she also happens to be aiming the arrow at the wood on which the bullseye is painted. However, there is an obvious sense in which the arrow aims at the wood only indirectly, because of the arrow aiming at the bullseye that is painted upon it. That the complex whole C10 is good for its own sake entails that it is fitting to favor C10 for its own

---

250 It may be remembered from previous chapters that Rønnow-Rasmussen says something very similar in his discussions of personal value. He asks how the features of a person can possibly figure in the intentional content of our love, say, without us also having to say that the attitude is likewise directed toward said features. His answer is simply that features may figure in the intentional content of our love as mere *identifiers* (2008, pp. 498–499). This means that they are not necessarily the primary objects of our love or even the features we would point to if asked why we have the attitude in the first place. If they were, then they would be more suitably regarded as *justifiers* (2008, p. 498). The point is that although features may figure in the intentional content of our love as both identifiers and justifiers, they certainly do not have to. Perhaps it is the case that when we favor a complex whole *simpliciter* while only favoring its parts *per accidens*, those parts figure in the intentional content of our attitude as identifiers.
sake. However, since C10 has a part that is bad for its own sake, it is fitting to favor C10 for its own sake only in such a way that the bad part is favored *per accidens*. Once this condition is met, there is an important sense in which the state of affairs that Mary is in pain is not the object of a positive favoring.

To further illustrate some of these points, suppose once more that we are considering something as grand as a possible world, where every part happens to be essential. We have seen that there can be reasons to favor possible worlds without there being reasons to favor all their respective parts. If what has been argued is correct, then it is just as reasonable to suppose that we can favor possible worlds without thereby favoring all their respective parts. This seems like the right result. To favor the possible world in which we are living is not to favor everything bad that ever happens within it. It is to direct an attitude toward the possible world in which its parts, be they terribly good or terribly bad, are only figuring in the attitude *per accidens*. The same line of reasoning appears reasonable when it comes to the favoring of individual persons. My love for a specific human being must not transmit to all her negative aspects. Those negative aspects are perhaps represented in my love in some sense, but this is not the same sense in which she is herself represented in my love.

Placing emphasis on discerning attitudes in this way can help defuse Olson’s objection, and it also teaches us something important about the connection between value and reasons for attitudes, but it should be mentioned that this is not the only strategy available to proponents of organicism and the principle of organic unities. Another way of diffusing the objection is by placing emphasis on the normative end of these matters. When an object has final value, then it is fitting to either favor or disfavor the object or its own sake. The relevant notion of fittingness is to be understood in terms of reasons. However, as we noted in the introductory chapter, the reasons in question need not be moral and they need not be so strong to amount to any form of rational requirement or ought. In short, the reasons that are of relevance to the fitting-attitudes analysis are justificatory but also of the *pro-tanto* variety. Their purpose is to call upon us to have certain responses, to recommend certain attitudes or acts, but not to make these responses obligatory. Once this is kept firmly in mind, the suggestion that we can have reasons to favor complex wholes whose parts we may have reasons to disfavor very strongly, should also become more palatable. The balance of reasons may even come out in such a way that the parts ought not to figure in positive attitudes even *per accidens*. Let us move on now to the following comment by Brännmark, which closely relates to what has just been said:
This would mean that although, e.g., sadistic pleasure is bad *qua* sadistic pleasure, it would still have some positive value simply *qua* pleasure. This is clearly a possible position, but a more reasonable one is that we cannot really view this instance of pleasure both as simple pleasure and as sadistic pleasure without suffering from some kind of normative schizophrenia: if a certain instance of pleasure is sadistic, then the element of sadism spoils it entirely (2001, p. 227).

This objection is independently endorsed by both Olson and Zimmerman. The former states that the theory of organic unities “seems to evoke a form of ‘evaluative schizophrenia’ in the evaluator” (2004, p. 42). With this objection, Brännmark wants to show that organicism commits us to the possibility of there being reasons to hold very different kinds of attitudes toward the same complex whole. The reason is that complex wholes can be good in virtue of some parts and bad in virtue of others, or good relative to some of its aspects and bad relative to some others. Against this line of reasoning, we can repeat some of the responses already mentioned. The practice of having distinct kinds of attitudes toward the same object is not as problematic when we notice that the attitudes are themselves concerned with distinct aspects of the object in question. The attitudes are fitting only to the extent that the object is represented within the attitudes in a way that is appropriate and compatible with the other attitudes we may have reasons to adopt. If we add to this the observation that the attitudes in question are called for by reasons of the *pro-tanto* variety, then Brännmark’s worry also seems to me to become somewhat less worrisome.

Finally, that there can be reasons to have different sorts of attitudes toward the same objects is guaranteed so long as we admit that objects need not be completely good or bad. Interestingly, Brown (2007, pp. 459–460, 461–463) makes a similar point and uses it in opposition to the theory of conditionalism. He suggests that if objects can lose or acquire final value depending on their place in different contexts, then the very same objects can be simultaneously good or bad relative to those contexts. Perhaps there is something to this. A painting could have final value as a part of a person’s awareness, a wall, a decor, a house, a neighborhood, a town, an aesthetic tradition, the cultural heritage of particular group of people, a possible world, and so on. Some of the resulting final values may be of the positive kind and others may be negative. We may have reasons to direct both positive and negative attitudes toward the very same objects—to welcome their contribution to certain complex wholes and to lament their contribution to others. No matter how banal an issue it might seem
at first sight, whether we should favor hanging up a painting on a wall can turn out to be a very complicated matter indeed.

Brown (2007, pp. 461–462) also argues that conditionalism is really the theory that risks severing the link between value and reasons. At one point during his discussions, Brown asks us to imagine a philosophy professor who is about to order food at a restaurant. The professor is of a mind to buy a plate of chips and considers whether she should get them with gravy or without. The reason for her indecision is that while experience has taught her that the complex whole made up of chips and gravy is better than chips alone, it has taught her very little about the value possessed by chips and gravy as parts. To the waiters waiting for her to decide on her order this will of course seem absurd. “So far as his evaluative reasons are concerned,” Brown insists, “our hapless customer has nothing left to decide. The values of objects as mere parts, about which he is undecided, have absolutely no bearing on whether he should get the gravy” (2007, p. 462). However, this seems to indicate that the value possessed by parts, in this case at least, matters very little from the perspective of reasons. “Our evaluative reasons,” Brown maintains, “are fully explained by the values of objects as wholes” (2007, p. 462).

The point is that the weightiest reasons we have for action tend to be given by their corresponding complex wholes. Observations about the potential profligacy of final value under conditionalism are meant to illustrate the point. To my mind, this argument is not very convincing, because it fails to appreciate the nature of the reasons that are given by the relevant parts. The values brought by the parts are not as inconsequential as Brown suggests, but the reasons that the values correspond to end up not being the decisive ones. This does not mean that the reasons in question are completely disabled and that they stop being reasons. My view is that neither organicism nor conditionalism incorporate failsafes against the sort of structure and behavior that might negatively affect decision making—or that would, at any rate, make our first-order theories about value that much more complicated. On the other hand, perhaps we should not expect them to. The sort of profligacy that has just been discussed is perhaps more appropriately countered by formulating a narrower picture of which objects have final value and which do not. We return to this issue in a moment. In the next section, I look closer at some of the first-order issues that are invited by these sorts of discussions and consider the possibility of accepting a hybrid view about the structure and behavior of final value. To reiterate, this hybrid view leaves the door open both for the theory of conditionalism and organicism.
On the Prosperity of the Sinful and the Possibility of Hybrid Views

Lemos (1994, p. 43) argues that conditionalism misses what is so repugnant about apparent cases of undeserved happiness. What makes such cases repugnant, he says, is that a person has a value that he does not deserve. Yet, according to conditionalism, it is rather that the happiness suddenly lacks a positive value. Lemos anticipates the reply that what is upsetting is rather that a person has something which is good for him, but that this is not enough to show that it is finally good. Here he appeals to the theoretical economy of organicism. The picture it affords of these kinds of cases does not have to involve the concept of personal value alongside that of final value. Even if the concept of personal value makes sense, then, a theory which does not have to invoke it to explain what is going on in these specific kinds of cases would be preferable. If he was aware of it, Lemos could add to this argument by invoking the previously demonstrated amenability of personal value to the distinction between final and instrumental value. After all, it seems like a reasonable suggestion that pleasure, in so far as it is good for a person, is also good for its own sake. Be that as it may, Olson (2004, pp. 45–46) responds to the worry raised by Lemos by avoiding having to make direct reference to the value that happiness has. The fact that the sinful person experiences happiness is not necessarily repugnant because happiness is good for its own sake or even good for the person. Roughly, Olson suggests that it is enough to notice that happiness feels rather nice and that we do not want the sinful person to have such nice experiences.

Hurka (1998, p. 311) provides an objection that is like the one raised by Lemos, but which focuses on the other side of the coin. He considers the fact that a sinful person is being subjected to pain. We are to suppose that the fact in question is a result of the sinful person being punished for the sins she has committed. Hurka thinks that it is bad to be sinful and that it is bad to be subjected to pain. The relevant good that results from their combination is to be located within the complex whole. What this means is that there are reasons to favor, in a somber way, the complex whole that a sinful person is subjected to pain, while there are reasons to lament the infliction of pain as such. It is very interesting that Hurka here explicitly embraces the idea of ambivalences that Olson finds so strange and argues that there are cases where such ambivalences are fitting. Proponents of conditionalism would instead assign the relevant good to the infliction of pain. It is here we will find the value that results from its infliction upon a sinful person. Hurka does not think that this is plausible, for retribution calls upon somber attitudes on our part, in a way that it would not if the infliction of pain was merely good. Not surprisingly, Olson (2004, p. 479) is
not impressed. He thinks that Hurka underestimates the variety and sophistication of the attitudes that may be called for by objects that are good. This is true enough. Just because the infliction of pain is good in this case, does not mean that there are reasons to feel pleasure:

An appropriate response toward cases of retributive punishment and deserved pain reasonably consists in an honoring of the fact that deserved pain is being inflicted, i.e. that amends are being made, and it is equally reasonable to suppose that this attitude ought to be characterized by a ‘sombre’ feeling tone. (2004, p. 47).

I shall not go into further details about this debate. It suffices to say that when it comes to the kinds of examples we have just looked at, my intuitions, and hence my first-order sympathies, tend to lie with Lemos and Hurka. The fact that the happiness of the sinful person is objectionable to me is because the sinful person would then have something that is good for its own sake.\textsuperscript{251} To my mind, Lemos is right to suggest that this is something that she does not deserve. Olson tries to explain this away by saying that it is enough that the sinful person would have something she would desire or enjoy. To explain why the case is objectionable, then, we need not judge happiness to be good in any literal sense. This seems wrong to me. There may be many things that the person in question would desire and enjoy which I would not begrudge her even though she had sinned. Were it not for the fact that it is good for its own sake, happiness would perhaps be among those things. Suppose the sinful person had an intense desire to be punished for the sins she has committed and that this would even cause her to feel enjoyment—to have an experience of value. Setting aside any thoughts of the happiness it might also cause her, this is a desire I would probably not mind see fulfilled in this case. Admittedly, retribution is not an ideal that I find generally appealing, but in so far as it makes sense, it seems to me not to be about depriving people of things that they happen to desire or enjoy, as much as it is about depriving them of values that they do not deserve. I take it that this is one of the ways in which retribution differs from mere vengeance.

When it comes to the case of the sinful person being subjected to pain, it seems equally plausible to me that the pain as such should be met with negative

\textsuperscript{251} Obviously, this is objectionable to me provided that the sins that have been committed are sufficiently severe, the sinful person has not shown any remorse, the happiness she experiences is of a sufficiently high intensity, and the happiness is experienced too soon after the sins in question have been committed.
attitudes, even in cases when it is inflicted upon the sinful person. If there is a positive value that arises here at all, it must be located at the level of the complex whole. Now, we have seen that the arguments raised against organicism in the previous sections were not particularly strong. Although they appeal to some of my intuitions and first-order convictions, perhaps the arguments just raised against conditionalism are not very strong either. It would be very surprising if there was an uncontroversial first-order example that could move any one of us toward one theory rather than another. In the end, much of this discussion becomes a mere table-thumping over first-order intuition. It would be less surprising, on the other hand, to find that all the theories mentioned in this chapter capture ways of talking and thinking about final value that are common to everyday life. If my arguments from the preceding sections are right, then from a second-order standpoint, aspects of both organicism and conditionalism appear equally valid. Of course, all this invites questions about the extent to which disagreements concerning the structure and behavior of value are likely to be settled in the second-order domain to begin with. If my arguments in previous sections are right, then the connection between value and reasons for attitudes do not seem very helpful in this regard.

The theories brought up in this chapter represent distinct ways in which final value can be related to other values at separate levels of the structure of the world. The arguments raised against the theories in this and the previous sections are meant to show that one of them ought to be rejected because it would exclude ways of talking and thinking about value that fits well with our first-order convictions. However, if this understanding of the debate is right, and much of the discussion ultimately turns out to be a first-order one rather than a second-order one, then one may also begin to wonder why the debate looks the way it does. If we are to discuss issues concerning the structure and behavior of final value from the perspective of the first-order domain, then perhaps it would be more helpful to start out from specific first-order theories. Suppose we accept classical hedonism and think that pleasure is the only thing that is good for its own sake and that its peculiar subjective qualities make it so. We also accept evaluative atomism as a first-order theory about the structure and behavior of value. Then the examples that are meant to show that final value is subject to holistic behaviors will of course start to seem like question-begging attempts at showing hedonism to be wrong.

It should also be pointed out that none of the objections we have so far looked at will retain much of a punch if we do not assume, as seems to be the norm in these debates, that organicism and conditionalism are mutually exclusive in the sense that they are perfectly general theories about final value.
Instead, we might suppose that both organicism and conditionalism represent legitimate ways of talking and thinking about final value. These distinct ways of thinking and talking about final value may not be compatible with respect to the same final values, but that is not surprising: Some final values might be intrinsic and so behave in the manner suggested by organicism, while other final values are extrinsic and so behave in the manner suggested by conditionalism, or even radical holism. Hurka (1998) recognizes this, as he contends that even though conditionalism is a reasonable stance, there are certain values that he prefers to interpret through the framework of organicism. McKeever and Ridge (2013, pp. 280–281) offer a more explicit defense of this sort of view, by arguing that the combination of organicism with other theories about the structure and behavior of final value is itself an organic unity in the sense that Moore had in mind. In this chapter, rather than trying to establish that one of the two theories is correct, I have tried to clarify what they commit us to and what they do not commit us to. In the end, I do not think that either theory must be generally correct, but that both represent understandings that are both conceptually and substantively feasible.

Concluding Remarks

Contemporary discussions concerning the structure and behavior of value have often centered around two distinct questions. Evaluative holism can be understood as an attempt to answer these questions in a way that avoid the troubles that afflict evaluative atomism. More specifically, evaluative holism is meant to be better at accommodating apparent holistic behaviors in final value. Thus understood, evaluative holism can come in several different varieties, two of which have been the focus of this chapter. Organicism claims that the final value of complex wholes cannot be analyzed as a function of their respective parts, but that the values of those parts are retained irrespective of the context of their occurrence. Conditionalism, on the other hand, claims that parts may lose or acquire different final values depending on what complex wholes they are parts of, but that the final value of the complex wholes themselves can still be analyzed as a function of the values of their respective parts. After drawing up the field and explaining some of the theoretical motivations and considerations that underlie the theories in question, efforts were made to explain their respective mechanisms.
Conditionalism follows straightforwardly from some of the assumptions supporting this work. One such assumption is that final value can have both a dependence base and a resultance base incorporating extrinsic properties. This means that when an object is good or bad for its own sake, it may be so in virtue of, or as a result of, properties that go beyond the intrinsic nature of the object. It is an implication of this view that parts could lose or acquire new final values depending on the complex wholes they are parts of. Organicism did not seem to follow quite as straightforwardly from the assumptions supporting this work, but it seems to be perfectly compatible with them. The idea is that the attitudes we have reasons to direct toward complex wholes need not match both the polarity and the intensity associated with the sum of the attitudes that there are reasons to direct toward their respective parts. It is an implication of this view that the final value which accrues to complex wholes can also fail to match the value that accrues to their respective parts. With this understanding in place, the question was asked which of the two theories can capture the apparent holistic behaviors of final value and offer escape from evaluative atomism.

Dancy argues that the notion of valuable complex wholes having neutral parts involves a breach of the connection between values and reasons for attitudes. If what I have said is right, then this argument fails for several distinct reasons. There is no incoherency to be found in the claim that complex wholes with final value may be constituted by worthless parts. There are even doubts as to whether there must be reasons to favor the essential parts of complex wholes that are valuable for their own sakes. Stratton-Lake, Brännmark, Olson, and Zimmerman raise the related worry that proponents of organicism would commit themselves to there being fitting ambivalences and normative asymmetries that there are strong reasons to avoid. For example, we can have reasons to favor complex wholes even when we have no reasons to favor their respective parts, we can have reasons to favor complex wholes even though we have reasons not to favor their respective parts, and we can have reasons to favor complex wholes with respect to certain parts but not with respect to others. It was shown that the unpalatability of this result can be mitigated by focusing on the kinds of attitudes that it is fitting to have and to remind ourselves of what their being fitting consists in.

The chapter was concluded by some general thoughts about the extent to which disagreements about organicism and conditionalism can be settled in the formal arena. Here the point was made that once discussions about the structure and behavior of final value start to move in the direction of the first-order domain, then we would do well to conduct the theories in more explicitly substantive terms. If we employ narrow first-order theories that ascribe final
value only to one class of things, such as the experience of happiness, then many
of the arguments against the views discussed in this chapter will start to seem
like they are question-begging. If, on the other hand, we employ wider first-
order theories that ascribe final value to different classes of things, then the
notion that we must choose between the different views discussed in this chapter
will start to seem less tempting. A kind of pluralism, on which both organicism
and conditionalism will seem to represent feasible theories about the structure
and behavior of final value, might be our best option.
Chapter 7
Closing Remarks

The preceding work has been on the topic of final value, which is the normative property that is assigned to objects when they are judged to be either good or bad for their own sakes. I began to illustrate final value by contrasting it with instrumental value, which is the normative property that is assigned to objects when they are judged to be good or bad for the sake of their effects. This also helped me highlight the importance of final value and the central place that the concept appears to occupy in our evaluative talk and practices. I said that objects with final value are those that should be treated as ends in themselves, while objects with instrumental value are those that should be treated as means toward other things. Besides occupying a principal place in our evaluative talk and practices, final value was also claimed to be significant from the perspective of moral philosophy. I assumed that ordinary disagreements about how we ought to live, as well as systematic attempts to formulate ethical theories, are to be partly grounded in notions about what objects have final value. Over the course of the preceding work, final value was contrasted with many other types of value as well. For example, after the introduction, final value was contrasted with value in a way, which was understood in terms of the concepts of having value as a kind of thing and having value for someone.

More specifically, chapter two started out from an influential objection in value theory drawing its inspiration from the writings of Geach and Thomson. The objection states that the concept of final value is incoherent, since everything that has value must have value in a way. I tried to show that the objections fail and that the concept of final value turns out to be a rather fundamental one. This means that many other value concepts that are
represented in our evaluative talk and practices, including value in a way, can come in the final variety. The chapter was not only meant to defend the concept of final value from the objection and to contrast the concept with value in a way, but also to highlight the conceptual relations that obtain between the two. Indeed, most of this work is more accurately described as having been on the topic of final value and the relations that it stands in to other things. However, what I have had in mind in the chapters following the second one was not just conceptual relations like the one highlighted between final value and value in a way, but also more robust explanatory relations.

If an object is good or bad for its own sake, then the object must have other properties that explain why this is so. Behind every final value, there is a normative explanation. Another way of stating the point is to say that final value must depend upon other things. Chapter three attempted to give some measure of clarification of the nature of this dependence. In so doing, two related worries were meant to be allayed: One worry, suggested by the argument of Jackson, was that by recognizing that there is a strong connection between final value and nonnormative properties, we end up committing ourselves to, or at the very least inviting the suggestion of, an identity between the two properties; The other worry, suggested by the arguments of Hofweber, Daly, Koslicki, and Wilson, was that if there is no such identity, the notion of dependence ends up being suspect, lacking in cognitive content, or theoretically dispensable. The chapter first presented an argument to the effect that the connection between final value and nonnormative properties does not entail an identity. I then tried to show that the dependence of final value upon nonnormative properties is an important topic of study, although there are legitimate doubts about whether it is a real part of the glue of the universe.

Chapter two and chapter three comprised the first half of the book, which was largely, although not exclusively, defensive in character. The second half was somewhat more constructive and started out with chapter four, in which we took a closer look at some recent attempts by philosophers to expand upon our understanding of how final value can be affected by the context of its occurrence. The chapter took as its point of departure the writings of Jonathan Dancy, wherein attempts are made to argue that normative explanations can feature background conditions. Among other things, background conditions are meant to explain why objects end up being good or bad for their own sakes, but in a more peripheral manner than being the factors in virtue of which, or as a result of which, the objects have final value. The purpose of the chapter was to offer an explication of the concept of background conditions that serves as a simpler and faithful alternative to some of the more common accounts currently
figuring within the literature. Very roughly, the explication in question takes background conditions for final value to be the factors in virtue of which, or as a result of which, facts become reasons.

Aspects of the discussion in chapter three were carried over to chapter five, where the focus was moved toward the framework of constitutive grounds. The framework in question was developed by Włodek Rabinowicz (Rabinowicz & Österberg, 1996) and has been used to clarify what it might mean to say that final value depends upon our attitudes. In this way, the framework is also meant to clarify the disagreement between subjectivism, which insists that there is such a dependence, and objectivism, which denies that this is necessarily so. The argument was made that while the framework of constitutive grounds is promising, it needs an interpretation if it is to be as informative as we want it to be. The interpretation put forward in the chapter follows the general pattern set out by the explication of background conditions set out in chapter four. It states that constitutive grounds of value should also be understood in terms of the things in virtue of which, or as a result of which, facts become reasons for attitudes. According to the interpretation, then, subjectivism is the view that attitudes are responsible for affording facts their normative status, while objectivism is the view that denies that this is so.

In the sixth chapter of the book, we moved on and took a closer look at issues having to do with the relation between the final value of complex wholes and the value of their respective parts. I was specifically interested in discussing two different versions of evaluative holism: The first version states that the final value of complex wholes need not be equal or even proportional to the sum value of their respective parts; The second version states that the final value of an object can change or be lost altogether depending upon the context of its occurrence. Here the concepts of background conditions and constitutive grounds regain their relevance. Much of the chapter was dedicated to figuring out how we should understand each version of evaluative holism. The main point was not, then, to argue that any one version of evaluative holism was correct and that there are cases that can be used to prove it. The main point was rather to clarify what would be needed for each of the versions of evaluative holism to be true. Attempts were also made at one point to defend the first version of evaluative holism by answering some recent objections. The objections were formulated by Jonathan Dancy, Jonas Olson, and Michael J. Zimmerman and were meant to throw doubt on the possibility of complex wholes that are good for their own sakes and yet consist of only bad or worthless parts.

I wish now to say a few words about what the preceding discussions have ultimately come to—if anything. In his seminal writings within second-order
value theory, John L. Mackie (1977) famously argues that value is *queer* in the sense that it is utterly unlike anything else in the universe. In support of his argument, he brings our attention to several distinct aspects of value that he thinks are worrying: One aspect is that value seems to affect our behavior in a way that other things are not. Value has the property of being such that our knowing about it is enough for us to be disposed to act in certain ways.

If an object is good and we judge it to be so, then we will be motivated, at least to some degree, to protect, promote, or preserve it. If the object is bad and we judge it to be so, then we will be motivated, at least to some degree, to prevent, destroy, or suppress it. This makes value rather unique relative to other kinds of properties. By contrast, if an object has the property of being red and we judge it to be so, then we will not be disposed to act in any specific way toward the object. The property of being red becomes relevant from the point of view of our motivation only if we either like or dislike red things—or so the intuition goes.252 Another way of making the same point is to say that an awareness of value appears to be necessarily motivational in a way that an awareness of other things is not. The second and, in the current context, perhaps most relevant aspect of value that Mackie thinks is worrying is the one that he describes in the following oft-quoted passage concerning the connection that value stands in to other things:

Another way of bringing out this queerness is to ask, about anything that is supposed to have some objective moral quality, how this is linked with its natural features. What is the connection between the natural fact that an action is a piece of deliberate cruelty—say, causing pain just for fun—and the moral fact that it is wrong? […] The wrongness must somehow be ‘consequential’ or ‘supervenient’; it is wrong because it is a piece of deliberate cruelty. But just what *in the world* is signified by this ‘because’?” (ibid, p. 41)

---

252 There is also an epistemic dimension to the criticism of Mackie which I do not bring up here, but which has also raised a lot of attention within the literature. The idea is roughly that if value is as metaphysically strange as he suggests that it is, then the means by which we come to know about value must also be very special, perhaps utterly unlike any of the other epistemic faculties that we happen to possess. In other words, Mackie’s worry is that for value to be knowable to us, we must be endowed with an experiential capacity that is just as exotic and strange as the value it informs us of.
I read Mackie as saying that value is strange because its connection to the nonnormative is strange. One could respond to this line of reasoning by saying that the general phenomenon of dependence is ubiquitous and no less strange outside the normative domain than inside of it. However, I am not so interested in going down this route, although I suspect that there could be something to it. I will instead grant that there is a mystery regarding the dependence of final value upon other things that is in some respects special and unique. If this is right, then this work has been dedicated to relieving the dependence of final value upon other things of some of this special and unique mystery. This is not to say that the purpose of this work has been to vindicate evaluative realism, which is of course the main target of Mackie’s objection. He wanted to cast doubt on the suggestion that value is a genuine property or relation in the world. In my view, the endeavor of trying to get clearer about the dependence of final value upon other things has merit quite irrespective of whether we think that this suggestion is sound or not. For to understand the dependencies that are thought to connect final value to the rest of the world is also to come approach an understanding of the concept of final value.

Of course, I also make no claim to the effect that the dependence of final value upon other things has here been relieved of all its special and unique mystery. I think that this is impossible and so my aims have been considerably more modest than this. Among the many conclusions reached within the preceding work is the following: If we think that it makes sense to say that objects can have normative properties in virtue of, or as a result of, certain factors, then things like background conditions and constitutive grounds need not be so difficult to understand. There is at least one sense in which factors can play the role of background conditions and constitutive grounds in normative explanations that can be made sense of without using notions that we are not already committed to. In like manner, if the relationship of evaluative resultance works in the ways that I have so far described, then the dependence of final value upon other values, and the relationships that might hold between the final value of complex wholes and the values of their respective parts, should also not appear quite so foreign either. In addition to this, the relations between final value and value in a way, as well as derivative value and constitutive value, have been clarified and significantly elaborated upon.

Even if the arguments that I have made throughout this work are sound, it seems to me obvious that final value remains a very queer thing—queerer, perhaps, than even Mackie appreciated when he first presented his famous worries. For if am right about the dependence of final value, then it turns out to be a very thick relation, characterized by a crisscross of distinct factors affecting
how objects end up being good or bad for their own sakes. If I have done anything of worth here, aside from tightening our grip on certain aspects of the what it means for objects to have final value, then it is perhaps that I have cleared up a bit of the structure within broad normative explanations and highlighted where the queerness and mystery is to be found.
Appendix

On Circular Analyses of Value

Much of the preceding work was informed by certain key assumptions that were explicated and at least partly motivated in the introduction. If I were to hazard a guess, the most controversial of these key assumptions is the one involving the fitting-attitudes analysis. This analysis proceeds from the idea that final value can be defined in terms of the presence of reasons for attitudes. This means that when an object has final value, this is equivalent to there being reasons to hold positive or negative attitudes toward the object for its own sake. In the introduction, I pointed out some of the main benefits and difficulties associated with this analysis. I also endorsed a suggestion according to which the fitting-attitudes analysis could be made explicitly circular and have most of its benefits, but fewer of its difficulties, preserved. I wish to elaborate on some of these topics in the following discussion and offer an additional defense of the suggestion.

Philosophers tend to treat the fact that a conceptual analysis is circular as a strong reason to reject it (Burgess, 2008, p. 215; Keefe, 2002, p. 275). From a historical perspective, it is not difficult to see why this is so. For much of its life, analytic philosophy seemed to be just as characterized by an ideological commitment to reductionism as it was to any specific philosophical methodologies. Old habits die hard, and even long after the ideological commitment to reductionism began to weaken circularity was still seen as something to be avoided.253 It is mainly in the past few decades that things have

253 Quine (1951) was earlier described as suggesting that the concept of analyticity cannot be understood without reference to meaning, necessity, and synonymy, all of which appear in its
begun to change and that specific instances of circularity have been defended by philosophers. The following appendix aims to contribute toward this trend. The view it presents states that circular analyses are much like any other conceptual analyses in that they can be either good or bad. For the most part, it might be true that circular analyses tend to be bad, but the point is that their circularity is not necessarily what makes them bad. When circular analyses are bad, they are bad for partly the same reasons that other conceptual analyses tend to be bad, namely because they are either uninteresting or incorrect.

The following appendix anchors itself in the kinds of value theoretic examples that have been used throughout this book. My arguments will also be given support from some recent contributions on the topic of circularity that have been written from a more general methodological perspective (Humberstone, 1997; Keefe, 2002; Burgess, 2008). According to the fitting-attitudes analysis of value, evaluative concepts can be defined in terms of fitting attitudes. One question we might ask, then, is whether a circular version of this analysis can be both interesting and correct. The answer which will be provided says that the circular fitting-attitudes analysis is just as likely to be interesting as its noncircular counterpart and that it is even more likely to be correct. The reason

conceptual vicinity. This tempted Quine toward a rejection of the concept, but he was obviously not motivated by a concern for reductionism. Indeed, his critique of analyticity was partly seen as an attack on the reductionism championed by the likes of the Vienna Circle (as an aside, I find it quite curious how hostile Quine was to circular analyses, given his commitment to confirmation holism, which could be taken to involve a sort of justificatory circularity). Humberstone (1997, p. 249) also points out that circularity was often seen by philosophers as something to be avoided even before the advent of analytic philosophy and its initial commitment to reductionism. As an example, he mentions the critique that Butler raises against prominent accounts of personal identity and knowledge. Butler states: “And one should really think it self-evident that consciousness of personal identity presupposes, and therefore cannot constitute, personal identity; any more than knowledge, in any other case, can constitute truth, which it presupposes” (Butler, 1975, p.100).

An example is found in the previously referred to Dickie (1984). See also Keefe (2002, pp. 276–277), Burgess (2008, p. 216), as well as the overview in Humberstone (1997, pp. 256–259, 272–275), wherein additional sources and historical discussions can be found.

There is a far more boring line of critique with which we shall not concern ourselves with here. This line of critique suggests that while some circular analyses may be acceptable, they should not be referred to as ‘analyses’. The idea is that the term ‘analysis’ is meant to apply exclusively to reductive accounts specified in terms of necessary and sufficient conditions. This idea seems to me to appeal to a conceptual analysis of conceptual analysis that I do not agree with. Be that as it may, I also do not think it matters much what we happen to call circular analyses. The question I want to explore is more to do with their interest and applicability in philosophy and other theoretical disciplines. Should my continued use of ‘analysis’ prove too bothersome for some philosophers, they can mentally replace it with some other appropriate term that they happen to prefer, such as ‘explication’, ‘account’, or ‘story’.

254 An example is found in the previously referred to Dickie (1984). See also Keefe (2002, pp. 276–277), Burgess (2008, p. 216), as well as the overview in Humberstone (1997, pp. 256–259, 272–275), wherein additional sources and historical discussions can be found.

255 There is a far more boring line of critique with which we shall not concern ourselves with here. This line of critique suggests that while some circular analyses may be acceptable, they should not be referred to as ‘analyses’. The idea is that the term ‘analysis’ is meant to apply exclusively to reductive accounts specified in terms of necessary and sufficient conditions. This idea seems to me to appeal to a conceptual analysis of conceptual analysis that I do not agree with. Be that as it may, I also do not think it matters much what we happen to call circular analyses. The question I want to explore is more to do with their interest and applicability in philosophy and other theoretical disciplines. Should my continued use of ‘analysis’ prove too bothersome for some philosophers, they can mentally replace it with some other appropriate term that they happen to prefer, such as ‘explication’, ‘account’, or ‘story’.
is that the circular fitting-attitudes analysis retains nearly all the theoretical benefits of its noncircular counterpart, while also managing to avoid some of its most difficult problems. Among the examples we will be discussing is the previously mentioned wrong-kind of reasons problem. To reiterate, the problem starts from the observation that there are conceivable cases where it is fitting to hold positive attitudes toward objects that are not good or where it is fitting to hold negative attitudes toward objects that are not bad.

Before we continue, I wish to mention a common worry in these contexts that has so far not been addressed in any of my discussions: On one understanding, concepts are irreducible and abstract entities, and on another understanding, they are mental representations. In the former case, conceptual analysis might seem to inherit the mystery of its target phenomenon. Concepts, in the spooky abstraction sense, do not exist, and so there is nothing there to really analyze. In the latter case, conceptual analysis, as it is traditionally practiced, might seem incapable, or at any rate inefficient, at capturing what really goes on in the heads of people. The latter objection is often formulated in terms of a challenge from cognitive psychology, which moves from the observation that ordinary concepts have been shown to often exhibit structures that cannot be captured by means of necessary and sufficient conditions. There is much that I would like to say in defense of conceptual analysis in the face of these objections, but which space does not permit me to mention. The following appendix will deal specifically with objections against circular analyses and it is meant to show that they merit as much thoughtful consideration as their noncircular counterparts. For now, I must unfortunately leave it up to readers to decide how much thoughtful consideration conceptual analyses are worth to begin with.

The Intuitive Case for Circular Benignity

Instead of saying that analyses can be either good or bad, let us say that analyses can be either acceptable or unacceptable. The question then becomes what reasons there are to suppose that circular analyses are generally unacceptable. Barring any issues having to do with the ideological commitment to

256 Furthermore, some of the comments I would like to make on this issue would be elaborations on the points already made by Sandin (2006). For the most famous formulation of the charge from cognitive psychology, to which Sandin aims most of his remarks, see Ramsey (1992).
reductionism, which most analytic philosophers ought to have given up on by now, the most obvious reason seems to be that circular analyses are either uninteresting or incorrect (Burgess, 2008, p. 215). To explicate these concepts in detail would be difficult, but perhaps we can make do with a rough and intuitive picture that may be elaborated upon and reconsidered in just a moment. For example, we might begin to say that analyses are interesting in the event that they represent the conceptual contents of their respective analysanda in a way that reveals a rich and theoretically useful structure. Otherwise the analyses are uninteresting. Analyses are correct in the event that they offer up sufficiently accurate representations of the conceptual contents of their respective analysanda. Otherwise the analyses are incorrect. With this rough and intuitive picture in hand, we shall move on to consider the issue whether all circular analyses can be generally rejected because they are uninteresting or incorrect. For the sake of illustration, we shall begin by considering the following caricatures:

**CR1:** The object x is finally good \(=_{df} \) x is finally good

**CR2:** The object x is finally bad \(=_{df} \) x is finally bad

**CR3:** The object x is finally good \(=_{df} \) moral philosophers at Oxford University tend to judge x to be good for its own sake

**CR4:** The object x is finally bad \(=_{df} \) moral philosophers at Oxford University tend to judge x to be bad for its own sake

I think that certain observations can be made about how interesting these analyses are, even without a general philosophical account of what it would mean for analyses to be interesting in the first place. One such observation is that CR1 and CR2 are not very interesting. They do not represent the contents of their respective analysanda in a way that reveals a rich and theoretically useful structure. This seems to me like a very good starting point. A general philosophical account of what it would mean for analyses to be interesting would have to be able to accommodate this observation to be considered plausible. Another observation along these lines is that CR3 and CR4 are very interesting indeed. They do represent the contents of their respective analysanda in a way that reveals a rich and theoretically useful structure. The structure is rich in that it involves such seemingly disparate concepts as that of value, judgment, tendency, and being a moral philosopher at Oxford University. The
way in which this is done is also theoretically useful, not least because it might
be of help in the formulation of some nifty ethical theories.\(^{257}\) In my view, a
general philosophical account of what it would mean for analyses to be
interesting would have to be able to accommodate this observation to be
considered plausible as well.

Because CR1 and CR2 are not interesting, but CR3 and CR4 are very
interesting, circular analyses cannot be generally rejected based solely on the
assumption that they are all uninteresting (cf. Keefe, 2002, pp. 277–278). For
any given circular analysis, it must be shown to be sufficiently uninteresting.
One way of doing so proceeds by arguing that the analysis fails to conform to
the requirements set out by a general account of what it means for analyses to be
interesting. Alternatively, if we have no such account on hand, we might
proceed by challenging the proponents of said analysis to demonstrate that it is
in fact interesting. If analytic philosophers, despite their best efforts, fail to
illustrate that their favored analyses reveal a rich and theoretically useful
structure, then we can at least take this to be a strong indication that the analyses
in question are not very interesting.

I think that certain observations can also be made about how correct these
analyses are, even without a general philosophical account of what it would
mean for analyses to be correct in the first place. One such observation is that
although they fail to be very interesting, CR1 and CR2 are undeniably correct.
Given the repetitive way in which they are formulated, they cannot help but
offer up sufficiently accurate representations of the conceptual contents of their
respective analysanda. Once again, this seems to be like a good starting point. A
general philosophical account of what it would mean for an analysis to be
correct would have to be able to accommodate this observation to be considered
plausible. Another observation along these lines is that although they succeed in
being very interesting, CR3 and CR4 appear undeniably incorrect. They do not
offer up sufficiently accurate representations of the conceptual contents of their
respective analysanda. After all, it is both conceivable and plausible that
philosophers at Oxford University are often wrong in their value judgments.
Indeed, it seems like a fair assessment that philosophers at Oxford University are
currently not in complete agreement about what is good or bad. Once again, I
submit that a general philosophical account of what it would mean for an
analysis to be correct would have to be able to accommodate this observation to
be plausible as well.

\(^{257}\) For example, one such theory states that we ought to live in such a way that moral
philosophers at Oxford University would approve of our behavior.
Because CR1 and CR2 are correct, but CR3 and CR4 are incorrect, circular analyses cannot be generally rejected based solely on the assumption that they are all incorrect. In this context, no general account of what it means for an analysis to be correct will be as helpful as in the previous case. For any given circular analysis, it must instead be shown to be incorrect in whatever ways we usually show that analyses are incorrect, be they circular or not. As far as I am aware, showing that analyses are incorrect is done by several methods, one of which involves subjecting them to counterexamples along the lines of the above. Once such counterexamples have been presented that seem relevant and important to the case at hand, we can take this to be a strong indication that the relevant analysis we are trying to evaluate is not correct.

Two caricature examples of circular analyses have just been used for illustration. One of the examples turned out to be correct but uninteresting and the other example turned out to be interesting but incorrect. For any given analysis, we want analyses that are both interesting and correct. Although it is a bit unclear how we should balance the respective disadvantages here, it seems reasonable to suggest that we have equally strong reasons to reject both caricature examples. We just need to remember that to reject the examples we do not need to think that both fail to offer up correct representations of the conceptual contents of their respective analyses. To be unacceptable and to be incorrect are of course different things. With this kept firmly in mind, we shall now move on to consider the account that was briefly mentioned in the introductory chapter, and which has been appealed to throughout the course of the book. The account in question is the fitting-attitudes analysis of value, which claims that value concepts can be analyzed in terms of fitting attitudes. To reiterate, the question we should try to answer is whether a circular version of the fitting-attitudes analysis could, in fact, end up being both interesting and correct. If the answer is that it can, then the account may also be acceptable. Let us remind ourselves briefly of what the applications of the fitting-attitudes analysis to the concept of final value looks like:

\[ F1: \quad x \text { is finally good} =_{df} \text { it is fitting to favor } x \text { for its own sake.} \]

\[ F2: \quad x \text { is finally bad} =_{df} \text { it is fitting to disfavor } x \text { for its own sake.} \]

\[ F3: \quad x \text { is finally better than } y \text { (} y \text { is finally worse than } x \text {)} =_{df} \text { it is fitting to prefer } x \text { over } y \text { for its own sake.} \]
\[ F4: \quad \text{x is finally equal to y (y is finally equal to x)} =_{df} \text{it is fitting to be indifferent between x and y.} \]

Here are the circular counterparts to F1–F4 that were developed in the introductory chapter:

\[ F5: \quad \text{x is finally good} =_{df} \text{it is fitting to direct an attitude toward x such that i) the fittingness of the attitude is grounded in factors that make x good for its own sake, and ii) the attitude represents the factors that ground its fittingness as the things that make x good for its own sake.} \]

\[ F6: \quad \text{x is finally bad} =_{df} \text{it is fitting to direct an attitude toward x such that i) the fittingness of the attitude is grounded in factors that make x bad for its own sake, and ii) the attitude represents the factors that ground its fittingness as the things that make x bad for its own sake.} \]

\[ F7: \quad \text{x is finally better than y (y is finally worse than x)} =_{df} \text{it is fitting to prefer x over y for its own sake so that i) the fittingness of the preference is grounded in factors that make x better than y (y worse than x) for its own sake, ii) the preference represents the factors that ground its fittingness as the things that make x better than y (y worse than x) for its own sake.} \]

\[ F8: \quad \text{x is finally equal to y (y is finally equal to x)} =_{df} \text{it is fitting to be indifferent between x and y, so that i) the fittingness of the indifference is grounded in factors that make x equally good as y (y equally good as x) for its own sake, ii) the indifference represents the factors that ground its fittingness as the things that make x equally as good as y (y equally good as x) for its own sake.} \]

Whether philosophers agree with the fitting-attitudes analysis or not, most of them would presumably acknowledge that F1–F4 would be interesting if they were correct. The structure revealed by F1–F4 involve such seemingly disparate concepts as value, fittingness, and attitude. That the structure thus revealed is also potentially useful from a theoretical perspective should be clear from just a cursory glance at the literature, not to mention the arguments and positions.
described in this book so far. Analyses like F1–F4 have been used in the literature for multifarious purposes such as i) analyzing a wide range of different value concepts which includes final value, instrumental value, kind-value, and personal value; ii) modeling a wide range of different value comparatives which includes more valuable than and equally as valuable as; iii) weakening some influential objections toward the notion that some objects can conceivably be just plain good or bad; iv) gaining leverage between different views on the bearers of value, including theories about whether concrete objects can also be judged to be good and bad; v) gaining leverage between different views about the structure and behavior of value, including theories about how the value of wholes relate to the value of their respective parts; and so on. Other examples can be cited, but these are enough to illustrate that F1–F4 have been thought to be very interesting indeed. The crucial question to be asked, then, is whether F5–F8 are any less interesting than F1–F4.

The suggestion might be made that we ought to deduct some points of interestingness because F5–F8 are reductive. Even if we are not ardent supporters of the ideology of reductionism, it seems intuitive to suggest that *ceteris paribus*, a reductive analysis is more interesting than one which is not reductive (Gertken & Kiesewetter, 2017, p. 3, n. 14). However, it is questionable whether all else really is equal in this case. After all, the analyses expressed by F5–F8 uncover structures that are several times richer than those which were uncovered by F1–F4. The former analyses do not just involve the concept of value, the concept of fittingness, and the concept of an attitude, they also involve a story about the normative role played by properties in normative explanations, as well as a story about how said properties are to be represented in the intentional contents of attitudes if these are to be fitting in the first place. It seems like a plausible rule of thumb that the more discerning the attitudes that are referred to by an analysis are, the more interesting the analysis will be. Philosophers need to be careful, then, because there might be a great deal of circular analyses that are actually more interesting than their noncircular counterparts. Of course, that F5–F8 are sufficiently interesting to warrant acceptance cannot be established conclusively here. Instead, I will content myself with the more modest point that F5–F8 do not seem so much less interesting than F1–F4 that only the former two would warrant rejection based on their not being interesting at all. In other words, philosophers who previously thought that F1–F4 were interesting should now find that the same is true of F5–F8.
Informativity and the Ban on Inferential Circularity

If what was said in the previous section is correct, then it seems reasonable to suggest that F5 and F6 do manage to represent the conceptual contents of their respective *analysanda* in a way that reveals a sufficiently rich and theoretically useful structure as well. Furthermore, even though they are circular, F5–F8 are not so much different from F1–F4 that it would be reasonable to suppose that the former cannot be just as theoretically useful as the latter. It has certainly not been shown that variations on F5–F8 could not also be used to illuminate the previously mentioned types of value concepts and comparatives. Nor has it been shown that variations on F5–F8 would be less capable of affording analytic philosophers with leverage in debates about the nature of value-bearers as well as the structure and behavior of the good and bad. Indeed, much of this book was written under the assumption that F1–F4 and F5–F8 are equal in terms of their theoretical usefulness, and it is at present difficult to discern any reasons for thinking that the assumption is mistaken. In this section, we shall start to consider these issues from a more general perspective. Let us consider the following quotation, which, among other things, outlines two different purposes that analyses might fulfill:

On the one hand, a definition might be suitable for imparting a term to somebody who had never encountered it at all. Perhaps we may call definitions that meet this requirement *pedagogically illuminating*. On the other hand, a definition might present a concept in such a way as to be philosophically enlightening to people who already understand the *definiendum*. It need not be suitable for imparting understanding to the neophyte, but it might display the concept in a way that reveals overlooked but illuminating aspects of a skill which competent speakers simply take for granted. I shall call a definition that fulfils this function *philosophically illuminating*. Although some definitions might be both pedagogically and philosophically illuminating, we have no reason to expect this in general, and absolutely no right to demand it of any would-be definition (Burgess, 2008, p. 217).

Rabinowicz & Rønnow-Rasmussen have made similar distinctions within the context of discussions about the wrong-kind of reasons problem. They maintain that circularities in the fitting-attitudes analysis ought to be tolerated if the
relevant analyses are illuminating in other ways than imparting conceptual
compentence: “The reason for providing an analysis in such cases is not so much
the need of criteria for a correct application of the concepts under
consideration,” they say, “but rather the importance of gaining a deeper
understanding of the conceptual framework” (2006, p. 120). Analogies can
perhaps be made here with the study of grammatical rules. The purpose of such
a study is not necessarily to learn to speak grammatically. A person who speaks
with perfect grammar might still be interested to know more of grammatical
rules. Her reason for doing so might be the satisfaction of her curiosity about
what is involved in being able to speak grammatically in the first place (cf.
Humberstone, 1997, pp. 255–256). The same general line of reasoning appears
to hold true for conceptual analysis. The purpose of engaging in such an
endeavor is not necessarily to attain conceptual competence in any strict sense,
but rather to uncover, in as simple terms as possible, the conditions that
underpin, or that can be abstracted from, the exercise of that competence.

The most prominent account of circular benignity that has ever been
formulated is without a doubt the one developed by Humberstone (1997). He
starts out by distinguishing between two distinct kinds of circularity, namely
analytic circularity and inferential circularity. Humberstone suggests that an
analysis displays analytic circularity if, and only if, the concept being analyzed is
either overtly or covertly used in the analysans of that analysis. We encounter an
overt circularity if, and only if, the analysandum of an analysis can also be found
explicitly mentioned in the analysans of that analysis. A covert circularity is
revealed when the analysandum of a given analysis can also be found explicitly
mentioned in the analysans of another analysis to which the former is linked in a
series of analytic steps. Suppose A is analyzed in terms of B, and B is analyzed in
terms of C, and C is analyzed in terms of A, then we have before us a case of
covert circularity. 258 An analysis suffers from the disadvantage of inferential
circularity if, and only if, “any argument or inference from premises claiming the
various conditions provided by the account to obtain, to the conclusion that the
concept applies, is itself circular in whatever sense an inference or argument can
be circular” (ibid, p. 250). Inferential circularity is obviously meant to be a
special case of analytic circularity. 259

258 All my examples have, for the sake of simplicity and concision, involved cases of overt
Circularity, but most of the observations made here are meant to apply across the board.

259 Humberstone provides an analysis of inferential circularity in terms of a certain sort of
procedural, or argumentative, circularity. This could be viewed as a major weakness, in so far
as we lack a firm grasp of what this procedural sort of circularity is meant to involve. Since it is
not my purpose to defend Humberstone’s account, we shall leave this aside for now.
Burgess (2008) is among the philosophers who largely accept the distinction proposed by Humberstone, but he formulates it in a way that is in certain respects more intuitive. He first points out that the *analysans* of a successful analysis will always express its *analysandum* either overtly or covertly. Burgess (ibid, p. 219) therefore suggests that an analysis is analytically circular if, and only if, its *analysans* is stated in terms of, and requires a prior grasp of, the *analysandum*. Regarding the explication of inferential circularity, Burgess makes similarly minor adjustment to Humberstone’s wording. Roughly, Burgess (ibid, p. 219) suggests that an analysis is inferentially circular if, and only if, the method of establishing the applicability of the *analysans* wholly or partly consists in the method of establishing the applicability of the *analysandum*. Return now to the distinction between analyses that are pedagogically illuminating and analyses that are philosophically illuminating. Humberstone (1997) seems to suggest that this distinction is very closely tied to his distinction between analytic circularity and inferential circularity. More specifically, I interpret him as saying that instances of analytic circularities that are also instances of inferential circularities can neither be pedagogically illuminating nor philosophically illuminating. Burgess largely agrees with this sentiment and states that an analysis that is both analytically circular and inferentially circular “could not meet the norm of informativeness—It could not teach us a word that was new to us, nor could it introduce us to, or illuminate, an application procedure—so it must be viciously circular” (2008, p. 221). To help illustrate this point, Humberstone (1997, pp. 262–268) uses an example from epistemology involving analyses formulated in terms of belief and knowledge. We shall consider Burgess (ibid, pp. 221–222) version of this example, since it is reminiscent of some of the caricatures mentioned earlier:

---

260 Other than refining the account of Humberstone by changing some of its wording, Burgess also provides his own version of that account that is meant to be more liberal. Burgess begins by distinguishing between two distinct kinds of inferential circularity: He suggests that an instance of inferential circularity will involve *full ungroundedness* if, and only if, “there is either no way of verifying or no way of falsifying the definiens save by proceeding via a loop” (ibid, pp. 220–221). The inferential circularity of an analysis will involve *partial ungroundedness* if, and only if, “either the verification or the falsification procedure contains a loop…” (ibid, p. 221). Burgess wants to allow for inferential circularities displaying partial ungroundedness, since he thinks that these can also be philosophically illuminating. He takes this to be the main difference between his version of the account under consideration and the one originally put forward by Humberstone. We shall take Burgess at his word here, which means that we do not have to consider his account in any closer detail. If FA5 and FA6 are allowed by the account formulated by Humberstone, which I shall soon argue that they are, then they will, *ipso facto*, be allowed by the version of this account defended by Burgess.
BC: $x$ is a cow $=_{df}$ Prince Charles knows that $x$ is a cow

The analysis expressed by BC is very much like those expressed by CR3 and CR4, which define value concepts in terms of what moral philosophers at Oxford University judge to be the case. CR3 and CR4 differ from BC in that the latter is a case of both analytic circularity and inferential circularity. To establish whether Prince Charles knows that $x$ is a cow, we must first establish that $x$ is in fact a cow. Things would be different if the analysis was formulated in terms of belief rather than knowledge (Humberstone, 1997, pp. 263–264; Burgess, 2008, p. 222). Part of the reason is that belief, unlike knowledge, involves no commitment to the veracity of its content. In other words, we could argue that Prince Charles believes that $x$ is a cow without assuming or having to involve any premises saying that $x$ is a cow. CR3 and CR4 appear to be on better footing than BC in this regard as well, for it seems equally possible to argue that moral philosophers at Oxford University judge that $x$ is good or bad, without having to assume or involve any premises to the effect that $x$ has any value whatever. Humberstone (1997, p. 263) would claim that what makes analyses like BC problematic is that the *analysandum* occurs in the context of the *analysans* in a way that is not suitably protected. By contrast, the *analysanda* of F5–F8 occur in the contexts of their respective *analysandum* in a way that does appear suitably protected. They are formulated in terms of attitudes of a sort that, much like beliefs or judgments, do not involve a commitment to the veracity of their respective contents. According to the account provided by Humberstone, then, the kind of circularity that is displayed by F5–F8 does not appear to be the kind that prohibits them from being philosophically illuminating. Now, consider next the following analysis offered up by Keefe (2002, p. 280):

**CK:** $x$ knows that $P$ $=_{df}$ $x$ has a true belief that $P$ which $x$ knows to be justified

The point of this analysis is to show that the account offered by Humberstone is too liberal in that it allows for a kind of circularity that invites infinite regress. In the case of CK, we do not have to assume or involve any premises to the effect that $x$ knows that $P$ to argue that $x$ knows that $P$, and so inferential circularity, at least in the sense specified by Humberstone, appears to be avoided. However, Keefe (ibid, p. 280) points out that we do have to ask whether $x$ knows that $Q$, where this is the proposition that $S$’s belief in $P$ is justified. The answer to this
question depends on whether $x$ has a true belief that $Q$ and knows that this belief is justified. It is not difficult to see that we are now on the precipice of an infinite regress. There are cases, then, where we will not be able to avoid a problematic sort of circularity, or where we will be led down the path toward infinite regress. The assumption at play here is that a circularity that involves an infinite regress cannot be philosophically illuminating. Keefe later discusses some potential solutions to this problem and criticizes them on different grounds, but we should not go into further detail here, as it is not really my purpose to defend Humberstone’s account in detail. For now, it is enough to notice that none of $F5–F8$ appear to represent the sorts of problem cases that Keefe worries about. We are therefore free to claim that, on at least one prominent account of circular benignity, which is still likely to apply to a great deal of cases with which we might be commonly met, the kind of circularity that is displayed by $F5–F8$ should not, in fact, be disallowed.

Certain analyses appear to me to constitute paradigmatic examples of circular benignity, just as certain other analyses appear to me to constitute paradigmatic examples of correctness. The caricature analyses expressed by $CIRC1$ and $CIRC2$ both appear to be correct and uninteresting, while the analyses expressed by $CIRC3$ and $CIRC4$ appear to be incorrect and interesting. More relevantly, philosophers commonly agree that $F1–F4$ would be interesting if they were true and, barring any ideological commitment that they might have to reductionism, the same should therefore hold for $F5–F8$. This line of reasoning takes care to deemphasize the need for a general account of circular benignity. This lends it a certain strength, in so far as we are skeptical of the prospects of such an account ever being formulated. Still, we might be interested in pursuing as general accounts as possible, in which case, I submit, the examples mentioned here

---

261 Keefe (2002) has argued that although he takes himself to provide one account of benign conceptual circularity, Humberstone really presents two closely related accounts. We have just finished considering one of them, but she identifies among his formal remarks on circularity a second and, in certain respects, narrower account. Her aim in the paper is to criticize both accounts by Humberstone and to argue that there can be no single general account of benign conceptual circularity. She claims that whether a circular analysis is acceptable is to do with context, and what the purposes and aims of the analysis are. The best we can do is therefore to be particularists and consider cases of conceptual circularity individually as we encounter them in our philosophical endeavors. I shall soon admit to the suspicion that Keefe is right about this (cf. also Blackburn, 1993b, pp. 259–279). However, I also wish to leave her more detailed critical comments about the account provided by Humberstone aside, since it is not my aim here to defend it. I shall also not be diving any deeper into the formal remarks made by Humberstone (1997, pp. 264–268, 275–280) about circularity, which try to account for the notion of a suitably protected context in terms of compositional independence. For problems with these formal remarks, see Keefe (2002, pp. 281–286) and Burgess (2008, pp. 224–225).
should work well as theoretical points of departure. This means that F5–F8 are not only given support by the fact that they conform to the requirements set out by the account of Humberstone, but, ironically, the intuitive plausibility of F5–F8 could also be taken to give some measure of support to that very account.

Wrong-Kind Reasons and the Virtues of Circularity

The next and final question to be considered is of course whether there are any reasons to think that the circular version of the fitting-attitudes analysis is incorrect. Here it is worth reminding ourselves that many philosophers have thought that, although it is very interesting, the standard applications of the fitting-attitudes analysis are incorrect. The most common reason for skepticism stems from the previously mentioned idea that fittingness can come apart from value. Not only are there conceivable cases where it can be fitting to hold positive attitudes toward objects that are not good, there are also conceivable cases where it can be fitting to hold negative attitudes toward objects that are not bad. This, in a nutshell, is wrong-kind of reasons problem. Many solutions to the wrong-kind of reasons problem have been proposed, some of which appear to me promising, if incomplete. It would not be possible to rehash them here and there is really no need to. The modest point I wish to make is that there are no reasons to think that the circular analyses expressed by F5–F8 get things more wrong than their noncircular counterparts in F1–F4. In fact, when we consider the wrong-kind of reasons problem specifically, the opposite appears to be the case. For the analyses expressed by F5–F8 are better equipped than F1–F4 to deal with the cases used to demonstrate the problem in question.

The reasons why F5–F8 are better equipped to deal with the wrong-kind of reasons problem were briefly explained in the introduction, but it is worth repeating them. When the Evil Demon forces us to admire the saucer of mud for its own sake, we are given reasons to do so since we want to avoid torture and death. According to the idea currently under consideration, the reasons in question are not of the right kind, because the property of the saucer of mud being such that favoring it for its own sake can help us avoid torture and death is not actually what makes it good for its own sake. When the Evil Demon forces us to admire him for his own sake, we are given reasons to do so since we want to avoid torture and death. The reasons in question are once again not of the right kind, because the property of the Evil Demon being such that favoring
him for his own sake can help us avoid torture and death is not actually what makes him good for its own sake.

In the introductory chapter to this book I also argued that there are other considerations, quite independent of the wrong-kind of reasons problem, why we should be open to the idea of introducing circularity in the fitting-attitudes analysis. If we not only allow ourselves such circularity in our understanding of the relevant reasons, but also in our understanding of the relevant attitudes, then we can more easily begin to clarify what it means to favor objects in distinct ways. I mentioned that this kind of approach has been adopted by several philosophers, but that the general notion that we cannot understand the attitudes that are relevant to the case of value without invoking evaluative concepts is often associated with Wiggins (1987). It is not unlikely that the future will see a revision of mental concepts that relieves them of their evaluative content, but as things currently stand Wiggins seems to me to have a point: The sense in which we currently admire things cannot be analyzed without reference to the apparent admirability of an object; love cannot be analyzed without reference to the lovability of its object; desire cannot be analyzed without reference to the desirability of its object; and so on. The kinds of non-doxtastic attitudes that seem relevant to value, then, cannot be analyzed without reference to some value notion or other.

If value is to be understood in terms of the presence of reasons for attitudes, then, it seems very unlikely that such an understanding could plausibly avoid conceptual circularity. In this appendix I have tried to defend the view that circular analyses are very much like any other conceptual analyses in that they can be either good or bad. In so far as circular analyses tend to be bad, it is not necessarily their circularity that makes them so. In fact, I suggest that circular analyses are bad for the same reasons that other conceptual analyses tend to be bad, namely because they are uninteresting or incorrect. Interestingly, Burgess (2008, pp. 230–233) suggests that circular analyses are often better than their noncircular counterparts and not only because of their ability to deal with the sorts of problems that we have just looked at. Far from being something that we ought to avoid, then, conceptual circularity can be a theoretical virtue. Burgess presents several reasons in favor of this view which revolve around the analysis of color phenomena, i.e. what it means for an object to be endowed with one color rather than another. For example, while thinking of what it means for an object to be red, Burgess suggests that “circular definitions can be the best available precisely because they gesture toward the need for supplementation in a way non-circular definitions do not”:
If our early philosophical counsellors were right, we need to experience redness to master the concept of redness. We need to experience redness in order to know what redness is. The dispositional theory of colour not only makes this need apparent, it locates the need in just the place where we would expect to find it—in our ability to know when something does, and when it does not, appear red. How could a definition that failed to do this be philosophically more perspicious than one that succeeded? (ibid, pp. 232–233)

Dispositional accounts of color suggest that objects are, say, red, if, and only if, competent observers experience them as being red under the right sort of circumstances (ibid, pp. 225–230). Burgess goes into significant detail as to how this is to be cashed out by offering up deeper stories about what competent observers are and how the right sort of circumstances should be understood. He thereby hopes to show that the resulting pattern of analysis involves analytic circularity without necessarily involving inferential circularity. This seems possible. Intuitively, we may argue that a competent observer is experiencing an object as being red under the right sort of circumstances without having to argue that the object is red. Now, to avoid being completely sidetracked into discussions about color phenomena, which are nearly as controversial as the proper understanding of value, we shall ignore most of the details of Burgess’s insightful discussion. The main point we should take away from the quotation above is the general one that there might be something about color that only a dispositional account of this sort could possibly capture. An account that lacked the circularity displayed here, and which thereby failed to emphasize the conceptual importance of having color experiences in the first place, could not hope to be quite as informative or correct as a noncircular counterpart.

If this is on the right lines, then perhaps a similar line of reasoning will apply, not generally to all circular analysis, but specifically to the kinds of circular fitting-attitudes analyses that are expressed by F5–F8. One way of understanding this would be to say that there are distinct aspects to the pull that value apparently exhibits. Value has a straightforwardly normative pull, which is explicitly captured by its formal connection to reasons for attitudes. Apart from this, perhaps value is also characterized by a distinct sort of evaluative pull, which cannot be completely couched in terms of the gravity of reasons. Among other things, this evaluative pull means that certain features of objects stand out and are, at least from the perspective of their observers, endowed with a certain shine. That it should be possible to convey this special shine by way of analysis
to someone wholly unfamiliar with evaluative concepts seems implausible. Rather, it must be experienced directly—in whatever sense it happens to be true that we have experiences of the good and bad. The circular fitting-attitudes analyses expressed by F5–F8 are more informative and more likely to be correct than their noncircular counterparts. This is not only true because they can help us deal with the problem of how to understand evaluatively relevant attitudes, or what characterizes reasons of the right kind. It is also true because F5–F8 are better able to accommodate the observations just made by leaving the shine of value as is. Although the analyses may not relieve value of all its peculiar sort of mystery, then, this should really be counted among their benefits, since, as I admitted earlier in my concluding remarks, value is inherently mysterious.
References


———. (2011). A New and Improved Supervenience Argument for Ethical


University Press.
York: Oxford University Press.
Bulletin, 107*(1), pp. 65–81
Manley, R. Wasserman (Eds.), *Metametaphysics: New Essays on the Foundations of
———. (1993). From Supervenience to Superdupervenience: Meeting the Demands of
Twin Earth. In: I. Ravenscroft (Ed.), *Minds, Ethics, and Conditionals: Themes from the
Clarendon Press.
(Ed.) and C. D. Broad (Transl.), Stockholm: Almqvist & Wiksell.
in M. Evens (Ed.), *Relations Models of the Lexicon*, Cambridge: Cambridge University
Sons.
Hooker, M. O. Little (Eds.), *Thinking About Reasons: Themes From the Philosophy of Jonathan Dancy*, pp. 265–285.


Quarterly, 56(222), pp.114–120.


