

EXPLAINING RIGHT AND WRONG

GEOFFREY H. FERRARI

ORIEL COLLEGE

TRINITY 2008



Submitted in partial fulfilment
of the requirements
for the degree of Doctor of Philosophy
at the University of Oxford.

Abstract

When an act is right or wrong, there may be an explanation why. Different moral theories recognize different moral facts and offer different explanations of them, but they offer no account of moral explanation itself. What, then, is its nature? This thesis seeks a systematic account of moral explanation within a framework of moral realism.

In Chapter 1, I develop a pluralist theory of explanation. I argue that there is a *prima facie* distinctive *normative* mode of explanation that is essential to moral theory.

In Chapter 2, I characterize normative explanation through its formal properties. I then draw on John Mackie's claim that moral explanations are queer to develop a powerful form of moral scepticism.

In Chapters 3–4, I reject attempts to reduce normative explanation to logical necessity, metaphysical necessity, or conceptual (analytic) necessity. The failure of these accounts is taken to reinforce Mackie's scepticism.

In Chapter 5, I defend a partial analysis of normative explanation in terms of irreducible normative laws. I argue that irreducible normative laws offer a realist, though non-naturalist, answer to Mackie's scepticism. The existence of irreducible normative laws then is defended as offering the best realist explanation of why rightness and wrongness supervene on descriptive properties.

In Appendix A, I discuss the claim the normative explanation has an essential connection to the motivation of virtuous agents. I defend this claim from certain difficulties posed by Jonathan Dancy's recent work.

Word Count (excluding abstract and bibliography): 72,500.

*For my father, Cataldo Ferrari,
and my grandfather, Roy Harrison Louch.*

Acknowledgements

In the preparation of this work, I have been helped by many others.

My chief philosophical debts are to my undergraduate tutor, Professor David Charles, and to my supervisor, Dr. Krister Bykvist. I thank David and Krister for all their support, criticism and encouragement in the long development of my ideas. I would also like to thank my peers in the philosophy department, especially Brian Ball, Wylie Breckenridge, Stephen Kearns, Ofra Magidor, Julia Markovits, Richard Price and Daniel Star for their friendship and for very helpful discussion of my frequently idiosyncratic views. Finally, I would like to thank the Fellows and staff of Oriel College, which has been my intellectual home for a decade.

On the personal side, I would like to thank my family and friends for their longstanding personal support. In particular, I thank James and Marion Bedford for always offering me a bed for the night. I thank Mark Blythe and Sander Nooij for their exceptional friendship and for raising the virtues of chappism to a heavenly sublime. I thank Lucie Cerna for being a wonderful friend and dance partner. I thank Axinja Hachfeld for being an inspirational friend and dancer. I also wish to thank my fellow grad students in Oriel College, especially Yuki, Caroline, Rhiannon, Philippa, Adrian and Farrah for so many wonderful times together. And to all my friends, thank you for being there.

I dedicate this work to my father, Cataldo Ferrari, and to my grandfather, Roy Harrison Louch, both of whom passed away during its preparation.

Contents

Contents	i
Introduction	vi
Outline of Thesis	viii
1 Explaining Right and Wrong	1
1.1 Explaining Right and Wrong	1
1.1.1 Explanations	5
1.1.2 General and Specific Explanatory Connectives	8
1.1.3 The General Problem of Explanation	14
1.1.4 A General Explanation Schema	16
1.2 Explanatory Pluralism	19
1.2.1 Causal Explanation	21
1.2.2 Inferential Explanations	22
1.2.3 Evidential Explanations	23
1.2.4 Conceptual and Metaphysical Explanations	25
1.2.5 Normative Explanation	27
1.3 The Nature of Normative Explanation	33
1.3.1 Normative Explanation and Normative Theories	38

1.3.2	Higher Order Normative Theories	44
1.3.3	Normative Explanation and Virtuous Agents	53
1.4	Conclusion	56
2	Normative Explanation	57
2.1	Introduction	57
2.1.1	Methodology	58
2.2	Formal Properties of Normative Explanation	60
2.2.1	Factiveness	61
2.2.2	Irreflexivity	62
2.2.3	Asymmetry	69
2.2.4	Transitivity and Intransitivity	71
2.2.5	Non-Monotonicity	74
2.2.6	Intensionality and Extensionality	75
2.3	Enablers and Disablers	81
2.3.1	Indicative Necessary Conditions	84
2.3.2	Counterfactual Necessary Conditions	85
2.4	Mackie's Question	87
2.4.1	Moral Dependence	88
2.4.2	Mackie's Argument from Explanatory Queerness	99
2.5	Analysing Normative Explanation	105
2.6	Conclusion	106
3	Inferential Accounts of Normative Explanation	108
3.1	Introduction	108
3.2	D-N Explanation	110

3.3	An Inferential Analysis of Normative Explanation	113
3.4	Against the D-N Sufficiency Thesis	115
3.4.1	Asymmetry	116
3.4.2	Irrelevance	118
3.5	Against the D-N Necessity Thesis	122
3.5.1	The Distortion Objection	123
3.5.2	Circularity Objection	126
3.5.3	The Internal Incoherence Objection	129
3.5.4	The Home Defeat Objection	132
3.6	Conclusion	134
4	Metaphysical and Conceptual Accounts of Normative	
	Explanation	135
4.1	Introduction	135
4.2	Metaphysicalism	136
4.3	The Metaphysical Necessitation Analysis	138
4.3.1	Against the Biconditional Law Sufficiency Thesis . . .	144
4.3.2	Against the Strict Conditional Necessity Thesis . . .	154
4.4	Supervenience Accounts of Normative Explanation	156
4.4.1	Formulating the Supervenience Analysis	161
4.4.2	Against the Supervenience Sufficiency Thesis	165
4.4.3	Against the Supervenience Necessity Thesis	171
4.5	Conceptualism	174
4.5.1	Direct Conceptualism	174
4.5.2	Indirect Conceptualism	178

4.6	Conclusion	187
5	Normative Accounts of Normative Explanation	190
5.1	Normative Necessity	192
5.2	Normative Reasons	196
5.3	A Nomological Analysis of Normative Explanation	205
5.3.1	Normative Laws	206
5.3.2	A Nomological Analysis of Non-Weighing Normative Explanation	210
5.3.3	From Normative Explanation to Deontic Supervenience	212
5.4	Conclusion To Thesis	222
	Appendices	224
A	Normative Explanation and Motivation	225
A.1	Introduction	225
A.2	Dancy on Action Explanation	227
A.3	Three Conceptions of <i>Motivating Reasons</i>	232
A.4	Three Arguments for <i>Success AP</i>	238
A.4.1	The Argument from Metaphysical Symmetry	238
A.4.2	The Argument from Explanatory Symmetry	242
A.4.3	The Argument from First Personal Intelligibility	247
A.5	Dancy's Argument For <i>Failure AP</i>	250
A.5.1	The Argument From Altered Motivation	252
A.5.2	The Argument from First Personal Accessibility	255

A.6 Conclusion	263
B A Formal Proof	265
Bibliography	267

Introduction

Most of us believe that there can be acts that are morally right or morally wrong. Most of us also believe that when an act is right or wrong, there may be an explanation why. For example, in March 2008 it was reported in the Times that Prime Minister Gordon Brown would allow some Catholic members of his Cabinet a free vote on the government's proposed legislation to permit research on human embryos. A senior Catholic clergyman, Cardinal Cormac Murphy-O'Connor, commented:

Certainly, there are some aspects of this bill on which I believe there ought to be a free vote, because Catholics and others will want to vote according to their conscience.¹

My overall aim in this thesis is to understand how explanations of right and wrong work. Different moral theories offer different explanations of the moral facts, but they offer no account of moral explanation itself. What, then, is its nature? And what is its place within moral theory? This thesis seeks a systematic answer.

An investigation into the nature of moral explanation is important for two reasons. First, it is highly plausible that every obligation *must* have an explanation. If one ought to return a book, having borrowed it, that may

¹The Times, 23rd March 2008.

be because one has implicitly promised to do so, or because it will be for the best, or perhaps because a virtuous person would do so. But, it seems absurd to suggest that one could be under an obligation for which there were no explanation at all. It would seem, then, that moral obligations *essentially* require explanations. That being so, any account of the nature of moral obligations is incomplete without an account of their explanation.

The second reason to investigate moral explanation concerns scepticism. John Mackie argued that a commitment to the objective truth of such moral judgements as ‘That is wrong because it is deliberately cruel’ presupposes an objective account of moral explanation. But, Mackie claimed, there is no objectivist account of moral explanation, no objectivist answer to the question of ‘just what *in the world* is signified by this ‘because’?’² Scepticism regarding *explanations* of right and wrong leads quickly to scepticism about right and wrong *themselves*, via what I call *Mackie’s argument from explanatory queerness*. It seems clear that objective moral facts require objective moral explanations. But, claims Mackie, there is no such thing as objective moral explanation. So, there can be, after all, no objective moral facts. Mackie’s argument presents an extraordinarily powerful challenge to moral objectivism that merits urgent attention.

A major aim of this thesis is to confront Mackie’s sceptical argument directly. I argue that some of the most credible attempts to account for objective explanations of right and wrong end in failure. But, I also argue that their failure shows the way to an objectivist account of moral explanation in terms of irreducible and objective normative laws. While this account

²Mackie 1977, p. 41.

is not without its difficulties, I argue that the moral objectivist need not despair of meeting Mackie's challenge on its own terms.

Outline of Thesis

In Chapter 1, I seek to place my discussion of explanation on a sound footing. I set out a general theory of explanation as propositional connection underpinned by necessity, and I argue that there are several *prima facie* distinct modes of explanation, underpinned by *prima facie* distinct varieties of necessity. This general theory of explanation is then applied to the moral case.

Explanations of moral facts, I argue, involve a distinctive *normative* mode of explanation, which I take to be the central object of my investigation. I introduce the idea of a normative mode of explanation by example, before developing two more theoretical characterisations. I identify normative explanation first and foremost through its central rôle in moral theory. Then, I offer a secondary characterisation of normative explanation through its connection with the motivation of virtuous agents. However, this secondary connection, which is also defended by *inter alia* Korsgaard and Stratton-Lake, faces a number of difficulties posed by Jonathan Dancy's recent work on the theory of motivation. Defending the connection between normative explanation and virtuous agents is important, but somewhat peripheral to my main line of argument, so I defer this work to an appendix. In Appendix A, I argue that Dancy's theory of motivation can be successfully challenged in a way that leaves the connection between virtuous agents

and normative explanation intact.

In Chapter 2, I consider the methodology appropriate to an investigation into the nature of normative explanation. I propose to begin with an examination of its *formal properties*. This methodology has the virtue of permitting a precise starting characterisation of normative explanation without presupposing any particular analysis of it. Having discussed a number of formal properties, I proceed to two further issues of importance. First, I defend the recent proposal, by Jonathan Dancy and others, to distinguish considerations that genuinely explain right and wrong from other considerations which function instead as *enablers* and *disablers*. I argue that every moral theory should make room for this distinction, and I sketch an account of enabling and disabling conditions which locates them at the level of the *second order* normative explanations. I then proceed to discuss Mackie's famous scepticism concerning normative explanation. I build on Mackie's work to provide a simple and powerful sceptical argument against both the possibility of objective normative explanation and the possibility of objective normative facts. Given its power and importance, the remainder of my investigation into the nature of normative explanation is structured as an investigation into the possibility of meeting Mackie's sceptical challenge.

The possible responses to Mackie's challenge fall into two broad categories, the reductive and the non-reductive. Given Chapter 1's assumption the explanatory connections must be underwritten by *necessity*, the natural candidates for a reductive analysis of normative explanation include the familiar forms of necessity, namely, logical, metaphysical and conceptual. Mackie himself stated that normative explanation 'cannot be an entailment,

a logical or semantic necessity³, but he offered no defence of those claims. Accordingly, in Chapters 3–4, I assess those candidate reductions in detail. My conclusion, however, is that Mackie was correct to claim that no reduction of normative explanation to logical, metaphysical or conceptual necessity is possible.

The failure of the major candidate reductions of normative explanation demonstrates the power of Mackie’s sceptical position. However, in Chapter 5, I argue that Mackie’s sceptical conclusion is by no means yet secure, because the possibility remains that normative explanation might be *irreducibly normative*. In my final chapter, I discuss three accounts of normative explanation in terms of normative *necessity*, normative *reasons* and normative *laws*. I argue that the normative necessity account faces fatal objection, and that the normative laws account is preferable to the normative reasons account because it can explain why rightness and wrongness supervene on descriptive properties.

³Mackie 1977, p. 41.

One

Explaining Right and Wrong

1.1 Explaining Right and Wrong

Moral facts often seem open to explanation. If one ought to return a book, having borrowed it, that may be because one has implicitly promised to do so, or because it will be for the best, or perhaps because that is what a virtuous person would do. Different first order moral theories recognize different moral facts and offer different explanations of them, but they offer no account of moral explanation itself. How, then, do moral explanations work? And what is their place within moral theory? This thesis seeks a systematic answer.

In this opening chapter, I set out the framework for my investigation into explanations of right and wrong. I begin by developing a framework for a pluralist theory of explanation. Next, I introduce the idea that there is a distinctive normative mode of explanation that is most salient for explanations of right and wrong. Finally, I argue that normative explanation is essential to the project of ethical theory itself, so that the aim of un-

derstanding it is well motivated. No such grandiose claims are needed, however, to begin my project. It is enough to think that some moral facts have explanations, and to be curious how they work.

I start with some preliminaries. It is natural to think of explanation as having both an epistemic and a non-epistemic component. It has an epistemic component through its connection with knowledge and understanding.¹ It has a non-epistemic component because knowledge and understanding are mental states whose truth or correctness depends on how things stand in reality. In this thesis, I aim to provide an account only of the *non-epistemic* element in explanation. I have little to say as regards the epistemic component.

By an *explanation*, I mean a proposition of a class that I will define in § 1.1.1. I take a proposition to be the *semantic value* of a token use of a meaningful declarative sentence. For the most part, however, my investigation will assume no particular account of the nature of propositions. Propositions might be abstract particulars, such as Russellian complexes, Fregean thoughts, or sets of possible worlds, or it may be that all talk of propositions is simply to be nominalized away. In two places, however, my arguments require the assumption that propositions are not sets of possible worlds. This claim is defended in § 2.2.6 and employed again in § 4.3.

Explanations, understood as propositions, are distinct both from sentences that express them and from token uses of those sentences. If there are acts of explaining which are not acts of expressing an explanation, such as pointing to a car to explain how one got here, my explanations are dis-

¹Cf. Salmon 1984; Strawson 1985.

tinct from those, too. Though I take explanations to be propositions, I will throughout use *sentences* to represent explanations and other propositions, speaking, for instance, of ‘the proposition that p ’. This sentential notation is consistent with the claim that propositions alone play the central rôle.

A speaker may successfully convey a complete explanation using an elliptical sentence, or by offering a partial explanation that contains only the information his audience does not yet comprehend as part of the complete explanation. My focus throughout this thesis is on complete explanations, expressed by non-elliptical sentences. I ignore partial explanations because they are subsidiary to complete explanations. I ignore elliptical sentences because either context provides whatever is needed for them to express whole or partial explanations, or else their pragmatic value is consistent with their being, one and all, literally false.

I will usually use *explanation* in a weak sense in which an explanation need not be true (or correct). In this sense, *explanation* is closer in meaning to *putative explanation*, but to speak always in that way would be needlessly long-winded. I will sometimes intend *explanation* in the stronger sense of a true (or correct) explanation, but context will make clear when that is so.

My investigation into the nature of moral explanation rests on two important starting assumptions. First, I assume a *cognitivist* picture according to which both moral judgements and moral explanations are truth-apt, and at least some of them are true. So, I will usually speak of *moral facts*, where by *facts* I mean simply *true propositions*. The relevant notion of a *moral proposition* is defined in § 1.2.5. Second, I assume an *objectivist* picture according to which the truth or falsity of a moral proposition is independent

of anyone's moral thought or experience.

My starting assumptions combine to yield a strong form of moral realism that is inconsistent with non-realist metaethical positions such as non-cognitivism and subjectivism. Non-realists should nevertheless find two sources of interest in my thesis. First, most contemporary metaethicists accept that it is important to stay broadly in contact with our ordinary, everyday moral intuitions, or what is sometimes called *folk morality*. Most contemporary metaethicists also accept that folk morality has at least a veneer of realism. For this reason, contemporary non-realists often seek to earn the right to say all, or nearly all, that the realist would want to say about both the content and nature of folk morality. Since much of my thesis consists of developments or refinements of apparently-realist folk morality, the onus is therefore on non-realists to show either that my conclusions can be recast in their own, non-realist terms, or else that they can be rejected. In § 2.4.2 I present one area in which non-realists may find this recasting of realist moral talk difficult. The second source of interest for non-realist metaethicists is John Mackie's sceptical attack on objective normative explanation. In Chapter 2, I argue that certain aspects of Mackie's argument have not received the attention they deserve, and that they provide the materials for a powerful sceptical argument against moral realism. This argument is of interest to realists and non-realists alike.

In this thesis, I will use the terms *moral* and *normative* interchangeably. Though the sphere of the normative is usually taken to be broader than that of the moral, encompassing everything from grammar to etiquette, I take morality to be part of the central core of important normative truths. For

my purposes, then, I can safely ignore the differences between them.

1.1.1 Explanations

By an explanation, I mean a proposition of a certain class. To identify the class of propositions that are explanations, I need to consider explanatory language. Explanations can be expressed with a variety of linguistic terms and phrases. Even limiting ourselves to moral examples, we can say that an action is morally obligatory *because* it was promised, that an action is wrong *in virtue of* causing unnecessary pain, that an action is permissible *due to the fact that* it causes no harm — and so on. I will not try to catalogue this wide variety. Instead, I divide our explanatory terms and phrases into the *propositional connectives*, and the rest.

A propositional connective is defined as follows. A function is a mapping from inputs, also called the *arguments*, to an output, also called the *value*. A *sentential* connective is a word or phrase in a language, which expresses a function that takes as its arguments two sentences, conceived of separately as the sentences p and q ², or together as the ordered pair $\langle p, q \rangle$, and yields as its value a compound sentence in which p and q appear as constituents. This is best illustrated by example. The English word ‘because’ can operate both as a sentential connective and as an adverb. The sentence

(1.1.1) He took his umbrella because of the rain.

illustrates the adverbial use. The sentence

²Throughout this thesis, I use the letters ‘ p ’, ‘ q ’, and ‘ r ’ either as sentence variables or as schematic letters whose substituends are sentences.

(1.1.2) He took his umbrella because it was raining.

illustrates the connective use. In its connective use, ‘because’ expresses a function that, when applied to the sentences ‘He took his umbrella’ and ‘It was raining’ yields as its output the compound sentence ‘He took his umbrella because it was raining’. It is helpful to represent sentential connectives using unbound sentence variables to indicate their argument places. So, the sentential connective expressed by ‘because’ is better represented as ‘ p because q ’.

A *propositional* connective is a sentential connective that also expresses a function from the propositions expressed by its input sentences to the proposition expressed by its output sentence. Put another way, a propositional connective is a sentential connective for which the proposition expressed by its output sentence is a function of, and hence entirely determined by, the propositions expressed by its input sentences. Consequently, when a sentential connective is also a propositional connective, we can effectively ignore its input and output *sentences* and focus only on the *propositions* they express.

I claim that explanatory sentential connectives such as ‘because’ are also propositional connectives.³ In any given context of use, the proposition

³A sentential connective is a two-place sentential operator. Many sentential operators are also propositional operators, but perhaps not all. For example, ‘Subject S desires that p ’ is a sentential operator, but the subject S may have an attitude of desire to the proposition that p when it is presented in the guise of the very sentence ‘ p ’, yet fail to have an attitude of desire to the proposition that p when differently presented. For example, Lois may desire that she marry Superman without desiring that she marry Clark Kent, even though, on many accounts, these desires have identical propositional contents. Consequently, the proposition expressed by ‘Subject S desires that p ’ may not be a function solely of the proposition that p . It may depend also on how the proposition that p is presented. I develop one of the consequences of this view in § 2.4.2.

expressed by the sentence ‘ p because q ’, is a function of the propositions that p and that q . Thus, the proposition expressed by a token use of the sentence ‘The Titanic sank because it [i.e. the Titanic] struck an iceberg’ would remain constant if the input sentences ‘The Titanic sank’ and ‘The Titanic struck an iceberg’ were replaced by other English sentences that expressed identical propositions in the context.

Not all explanatory terms and phrases operate as connectives. The phrase ‘in virtue of’, for example, often functions adverbially, as in ‘He is a father *in virtue of* having a daughter’. However, to many, perhaps even to all explanatory terms and phrases that can operate non-connectively, there correspond one or more related explanatory terms and phrases that can operate connectively. For example, to the explanatory non-connective phrase ‘in virtue of’ there correspond the connectives ‘in virtue of its being the case that’ and ‘in virtue of the fact that’. To ‘due to’ there correspond ‘due to its being the case that’ and ‘due to the fact that’. Furthermore, I will assume that every sentence in which an explanatory term or phrase is used non-connectively can be translated into one in which all the explanatory terms and phrases are used as connectives, without change in the propositional content expressed. For example, I claim that the proposition expressed by

(1.1.3) He is a father in virtue of having a daughter.

is identical to that expressed by

(1.1.4) He is a father in virtue of its being the case that he has a daughter.

In this way, the account of explanatory connectives can be extended to include non-connective explanatory phrases.

Let us say that an explanatory propositional connective such as ‘ p because q ’ expresses an *explanatory function*. I can now give two important definitions:

Explanation An *explanation* is the value of an explanatory function.⁴

Mode of Explanation A *mode of explanation* is an explanatory function.

I have now introduced the theory of propositional connectives and used it to define explanations and modes of explanation. I will now develop this theory in more detail.

1.1.2 General and Specific Explanatory Connectives

Explanations, I have said, are the values of explanatory functions expressed by certain propositional connectives. My aim now is to develop this idea further. Explanatory connectives exhibit wide variety, but again, I will not try to catalogue it. Instead, I divide explanatory connectives into the *general* and the *specific*. Some examples of general explanatory connectives are given in Table 1.1 on the following page.

I claim that each of the connectives listed in Table 1.1 expresses one and the same explanatory function in any given context of use.⁵ So, for example,

⁴Or, with apologies to Quine: to be an explanation is to be the value of an explanatory function.

⁵Here I overlook one technicality to do with the order of the arguments. For example, it may be that the connective ‘ p because q ’ maps the ordered pair of propositions $\langle p, q \rangle$ onto a proposition in which the propositions that p and that q appear in a given order iff the connective ‘That q explains that p ’ maps the inverse ordered pair $\langle q, p \rangle$ onto a

p because q
p in virtue of its being the case that q
p due to its being the case that that q
That p explains that q
That p makes it the case that q

Table 1.1: An Incomplete List of General Explanatory Connectives

I claim that in any given context the sentence ‘ p because q ’ expresses the same proposition as does ‘That q explains that p ’, for constant values of p and q . My position here is thus in agreement with, though more general, than that recently articulated by John Broome. Broome writes:

When I say one fact X explains another Y , I mean simply that Y obtains because of X . As I understand it, the explaining relation is merely the inverse of the because relation.⁶

There are two small differences between Broome’s view and my own. First Broome speaks of ‘*the* explaining relation’, as if there were only one. However, I will argue in § 1.2 that there is more than one mode of explanation. Second, Broome speaks of an explaining *relation*, not a *connective*.⁷ In recent unpublished work, Broome clarifies:

proposition in which the propositions that p and that q appear in the inverse order. If so, these connectives may express distinct, though inverse functions. Nothing important rides on this technicality, though, so I gloss over it.

⁶Broome 2004a, p. 32.

⁷What is the difference between a connective and a relation? A relation is a non-linguistic entity expressed by a *dyadic predicate*, which is an unsaturated phrase formed by removing two names from a meaningful declarative sentence in which those names occur extensionally i.e. in a referentially transparent context (Bostock 1997, pp. 74-5). A connective, by contrast, is a *dyadic propositional operator*. That is to say, it is a propositional function expressed by an unsaturated phrase formed by removing two or more constituent sentences from a meaningful declarative sentence (Bostock 1997, p. 18). The difference between a connective and a relation is thus to be understood in terms of the difference between a name and a whole declarative sentence.

What sorts of things are related by the explaining relation? Grammar allows various sorts of things to stand in this relation. For example, I said that natural selection explains evolution; there I was saying that one process explains another. But if we choose, we may harmlessly regiment our language by taking both explanans and explanandum to be facts. In the regimented language, we would say that the fact that natural selection occurs explains the fact that evolution occurs. Again: evolution occurs because natural selection occurs.⁸

Here I diverge slightly from Broome in taking the connective uses of ‘because’ and ‘explains’ to be primary. These technical details aside, my position accords with Broome’s.

I claimed above that the explanatory connectives listed in Table 1.1 express identical explanatory functions in any given context of use. But, this does not entail that they express identical explanatory functions across *different* contexts of use. In § 1.2, I defend the thesis of *explanatory pluralism*, according to which there is more than one explanatory function. Simply put, I defend the view that there is more than one way to explain something. So, I need to say something about the meaning of the general explanatory connectives given that there is more than one explanatory function they might express.

To discuss this question, let me introduce a piece of helpful notation. Henceforth, I will abbreviate the connective ‘That p explains that q ’ to the simpler ‘ p explains q ’. I will also allow adverbial modifications of this notation. So I will write e.g. ‘ p causally explains q ’ to abbreviate ‘That p causally explains that q ’. This notation requires a formal introduction because it departs from standard English.

⁸Broome 2008, p. 57.

I begin by noting that *general* explanatory connectives can be qualified in a variety of ways to yield a *specific* connective that indicates more precisely which explanatory function it expresses. For example, the connective ‘*p* explains *q*’ can be adverbially modified into the specific connective ‘*p* *causally* explains *q*’, which expresses the explanatory function appropriate to causal explanation. For grammatical reasons, some general explanatory connectives, such as ‘because’, do not permit this kind of adverbial modification. Here we might speak instead of a ‘causal use of “because”’, or we might indicate the relevant mode of explanation with a grammatical innovation, writing e.g. ‘The Titanic sank (causally) because it struck an iceberg’.

I propose that any account of the meaning of the general explanatory connectives must respect the following constraint. Take the list of specific connectives to be given by ‘*p* causally explains *q*’, ‘*p* inferentially explains *q*’, The point to observe is that a token use, in a context, of a general explanatory connective such as ‘*p* explains *q*’, expresses a truth iff either *p* causally explains *q*, or *p* inferentially explains *q*, or That is to say, a token use of a *general* explanatory connective such as *p* explains *q* in a given context expresses a truth iff there is some *specific* connection between *p* and *q*. Call this the claim of *Disjunctive Equivalence*. To deny Disjunctive Equivalence is to suppose that there might be such a thing as *bare explanation*, that is, that there might be a general explanatory connection which implies no specific connection. I know of no defence of this unattractive view, so I ignore it. The question now arises why Disjunctive Equivalence should be true.

Consider two possible explanations of Disjunctive Equivalence. The first view is that the general explanatory connectives stand to the specific explanatory connectives as *determinates* to *determinables*. This view promises to explain Disjunctive Equivalence as a general feature of the determinate/determinable relation. Disjunctive Equivalence holds for the same reason that an object is coloured iff it is red or it is green or it is blue or . . . , and so on for all the colours. If this view could be developed, it would also offer an account of what the various specific explanatory connections have in common. Just as red, green and blue are related as determinates of a single determinable, so causal and other specific explanatory connections would be related as determinates of a single determinable, *explanation*. However, if this account is to be properly developed, sense must be given to the idea that general explanatory connectives express an explanatory function that stands to the specific explanatory functions as determinate to determinables, and this, I claim, cannot be done. My rejection of this view rests on the assumption that determinates and their determinables must belong to distinct metaphysical categories. Just as the determinable *colour* cannot itself be a determinate colour, so, too, the determinable *explanation* could not be a determinate explanatory function. But, I assume that general explanatory connectives do express a determinate function (in a given context of use), in which case, they cannot express the determinable *explanation*. So, I reject the determinate/determinable account.

The alternative view is provided by *contextualism*. On this view, a general explanatory connective always expresses a specific explanatory function, but *which* function it expresses is determined by context. For example,

when causal explanations are most salient, uses of ‘*p* because *q*’ will likely express the causal explanatory function. But there are other, non-causal kinds of explanation, each with their own associated explanatory function, and in other contexts, a token use of ‘*p* because *q*’ may express other explanatory functions. Disjunctive Equivalence is, then, explained through the *context-sensitivity* of the general explanatory connectives. In any given context, a token use of a general connective will be contextually determined to express at least one, and in most cases, at most one, contextually salient explanatory function. We can add to this view the thought that the rôle of the *specific* explanatory connectives is to reduce the context dependence of a general connective. In effect, qualifying the general explanatory connectives, as it were, *decontextualises* them, with an effect roughly akin to disambiguation. In this way, the sentence ‘That *p* causally explains that *q*’ can express the explanatory function appropriate to causal explanations even in contexts where causal explanations are not most salient.

On the contextualist view, the difference between a *general* and a *specific* connective lies in their context dependence. The explanatory function expressed by general explanatory connective is determined by context, whereas the explanatory function expressed by a specific connective is determined by its non-contextual syntactic and semantic features.

In what follows, I will assume that the contextualist approach to explanatory connectives is correct. However, while contextualism avoids the problems that beset the determinate/determinable account, it offers as yet no answer to the question of how the various specific explanatory functions are related. What makes one propositional function *explanatory* and an-

other not? If there is more than one explanatory function, what unity is there among them? At least part of the answer, I suggest, is that all explanatory functions involve *necessity*. I develop this proposal in more detail in § 1.2.

1.1.3 The General Problem of Explanation

An explanation, I have argued, is the value of an explanatory function. A *general* explanatory connective expresses at least one, and usually at most one explanatory function in any given context of use. A *specific* explanatory connective non-contextually expresses just one explanatory function. *Explanatory pluralism*, to be discussed in § 1.2, may be roughly defined as the view that there is more than one explanatory function, *explanatory monism* as the view that there is just one explanatory function.

Let us take a moment to consider the explanatory functions as a class. Let the connective ‘*p* explains *q*’ be symbolised ‘ $[p * q]$ ’, and let us assume that which explanatory function this expresses is determined uniquely by context. One of the central questions to be faced by any investigation into the nature of explanation is to understand when and why explanations are true and false. Put another way, what is needed is a non-trivial account of their truth conditions. Given that explanations are the values of explanatory functions, an important question is how far the truth conditions of $[p * q]$ are determined by the truth values of its inputs. Is $[p * q]$ a total or even a partial truth-functor?

A *total truth functor* is a propositional connective that expresses a *total*

truth function, that is, it expresses a function from truth values to truth values that is defined for the entire range of inputs. When a connective is a total truth functor, it has the special property that the truth values of its output propositions are entirely determined by the truth values of its inputs. A partial truth-functor, by contrast, is a propositional functor which expresses only a partial truth function, one that is defined only for a limited range of inputs.

The standard view of explanatory connectives is that the truth of p and q is *necessary* but not *sufficient* for the truth of $[p * q]$. The truth of p and q is necessary for explanation because only truths explain, and only truths get explained. However, $[p * q]$ must be true for some true values of p and q , else no truths would be explained, but not for all, else every truth would explain every truth. The standard view is, thus, that explanatory connectives express the partial truth function illustrated in Table 1.2.

p	q	$[p * q]$
T	T	—
T	F	F
F	T	F
F	F	F

Table 1.2: Truth table for an arbitrary explanatory connective.

There is, then, a very general problem in the theory of explanation, that of giving a complete, non-trivial account of the truth conditions of explanations. Were explanatory functions total truth functions, we might say that their truth conditions consisted in nothing more than the relevant combina-

tions of truth and falsity of their constituent propositions. But, explanatory functions are not total truth functors, so a complete account of their truth conditions of explanations is still wanting.

1.1.4 A General Explanation Schema

In § 1.1.3, I posed a very general question about the truth conditions of explanations. Now I want to defend a very weak conception of how this question should be answered. Consider again our arbitrary explanatory function expressed by $[p * q]$, again reading this as ‘ p explains q ’. Roughly put, the idea is that, although the truth conditions of $[p * q]$ do not consist solely in the *truth conditions* of the constituent propositions expressed by p and q , they do consist solely in *something* to do with those propositions. I claim, that is, that the truth conditions of $[p * q]$ consist in some *relation* between the propositions expressed by p and q . Whatever this relation is, I will call it the *explaining relation*, symbolized as ‘ $E(p, q)$ ’. I propose, therefore, that a necessary condition of any successful account of the truth conditions of $[p * q]$ is that it should be expressible in the following schematic form:

Simple Explanation Schema (SES)

p explains q iff $E(p, q)$

SES provides the simplest expression of the framework that I will employ in this thesis, according to which the truth conditions of $[p * q]$ are to be given in terms of a relation between p and q . In some cases, however, it will be more perspicuous to employ a refinement of SES. The problem

with SES is that it may not always reveal enough structure. In particular, it is natural to permit explanations in which explanans, explanandum, or both, are *conjunctions*; but where that is so, it will be helpful to take the explaining relation to hold between *sets* of those conjuncts. Thus, letting \mathbb{P} be the set $\{p_1 \dots p_i\}$ and letting \mathbb{Q} be the set $\{q_1 \dots q_i\}$, my official policy is to employ the following refinement of SES:

Refined Explanation Schema (RES)

$p_1 \dots$ and p_i explains $q_1 \dots$ and q_i iff $E(\mathbb{P}, \mathbb{Q})$.

In RES, the explaining relation is re-conceived as relating *sets* of sentences that express propositions. However, for ease of presentation, I will drop the set notation whenever no ambiguity will result, and I will not decompose conjunctive propositions into their conjuncts except when necessary. So in practice, I will often end up using something that looks like SES.

The requirement that an account of explanation should be expressible in terms of my Refined Explanation Schema, RES, provides only a very weak constraint on a theory of explanation. In fact, there are two ways in which RES will be entirely trivial, or nearly so. First, there are as yet no constraints on the explaining relation, E , so nothing yet rules out that ‘ $E(p, q)$ ’ should be defined circularly as the relation between p and q such that p explains q . Second, suppose that $E(p, q)$ has some (possibly complex) truth condition C which, intuitively speaking, has nothing to do with p or q . Then, since I am happy to permit deeply extrinsic relations — analogous to so-called *Cambridge properties* — the explaining relation might yet be specified simply as the relation between p and q such that C . In these two

ways, then, RES can be either entirely trivial or nearly so. So, if RES is to express a substantive thesis, I must add two qualifications:

1. The explaining relation, E , must not be specified circularly in a way that depends on the concept of explanation itself;
2. The explaining relation, E , must not be deeply extrinsic.

My second qualification here is left deliberately vague, because I have no account to offer of the difference between deeply extrinsic or *Cambridge* relations and genuine relations. I stipulate only that the requirement that the explaining relation between p and q must not be deeply extrinsic does not imply that it must be *intrinsic*. RES would be a more interesting thesis if this claim could be defended. But, it cannot be ruled out at this stage that explanatory connections are largely extrinsic. RES is just strong enough to be substantive, but flexible enough to permit some extrinsicity.

I suggested earlier that part of what unifies explanatory functions is that they involve necessity. If that is so, then it is a necessary condition of p explaining q that it is in some sense *necessary* that if p then q . This view coheres well with the thought that explanatory functions are only partial truth-functions, because many varieties of necessity are not total truth functions. That is, for many interpretations of the necessity operator ' \Box ', the truth of $\Box r$ is not a total function of the truth of r : some truths are necessary, others are not. However, in most cases, $\Box r$ implies r , so $\neg r$ implies $\neg \Box r$. The thought that explanations involve necessity, vague though it is, may, therefore, help to explain the partial truth-functionality of explanatory functions.

1.2 Explanatory Pluralism

In § 1.1, I characterized the central aim of this thesis as understanding how explanations of right and wrong work. That statement of my aim is rough because there is, I claim, more than one mode of explanation for moral facts. In this section, I defend this claim, which I call *explanatory pluralism*, and describe a number of plausible alternative modes of explanation, the better to distinguish them from the *normative* mode of explanation that will eventually be my focus.

A mode of explanation, as defined above, is an explanatory function of the sort expressed by a token use in a context of an explanatory connective such as ‘*p* explains *q*’. Plurality in the modes of explanation is thus equivalent to plurality in the explanatory functions. There are, however, weaker and stronger versions of pluralism. *Superficial explanatory pluralism* claims only that there is more than one mode of explanation. The identity conditions of a mode of explanation, like those of all functions, are given by the identity of its mapping. So, it follows that if there are two modes of explanation, and two propositions expressed by the sentences *p* and *q*, such that *p* explains *q* in one mode but not in the other, then those modes of explanation are distinct and superficial pluralism is true.

Deep explanatory pluralism is, by contrast, the view there is more than one *fundamental* mode of explanation, in the sense that there is no mode of explanation to whose terms all genuine modes of explanation can be reduced. Whether a mode of explanation is fundamental is difficult to establish, because *reduction* is a difficult phenomenon to define. I will return

to the topic of deep pluralism and reduction in § 2.5. However, I note straightaway that superficial pluralism does not entail deep pluralism, because two superficially distinct explanatory modes might still be reducible to the terms of a single mode of explanation plus further conditions. Consider an illustrative case. Suppose we accept, as I will shortly argue, that among the genuine modes of explanation there are the *causal* and the *inferential*, which is the mode of explanation exhibited by valid arguments.⁹ Causal and inferential explanations are obviously superficially distinct, since not all valid arguments provide causal explanations. Still, causal explanations might be reducible to inferential explanations. For example, it might be that causal explanations are simply those inferential explanations which include statements of laws of nature among their premisses. This was, in essence, the view defended by Hempel and Oppenheim¹⁰ in their development of the *deductive-nomological* model of explanation. There has, of course, been a great deal of work to argue that causal explanation cannot be reduced to inferential explanation in this way.¹¹ However, there is an intuitive sense in which this project might have succeeded, and might thereby have shown that the causal and inferential modes of explanation were only superficially, and not deeply, distinct.

My aim in this section is to offer a brief defence of explanatory pluralism by outlining some plausible non-normative modes of explanation. In § 1.1.4, I suggested that it was a *necessary* condition of every explanation that it should involve a mode of *necessity*. One natural thought, then, is that

⁹Cf. Salmon 1984, p. ix.

¹⁰Hempel and Oppenheim 1948.

¹¹See e.g. Salmon 1984, esp. chs. 1–2 and 4.

different modes of explanation involve different modes of necessity. As Kit Fine writes:

A proposition is necessary if it *must* be true and possible if it *might* be true. On the face of it, there are different ways in which a proposition might be necessary.¹²

In line with this thought, I will associate some plausible non-normative modes of explanation with their associated modes of necessity. I will argue that the *superficial* distinctness of these modes of explanation is trite. I will then introduce the idea that there is a distinctive *normative* mode of explanation that is superficially distinct other modes. It is this normative mode of explanation that will eventually be my focus.

1.2.1 Causal Explanation

It is a natural view that many explanations, perhaps especially those within the natural sciences, are causal.¹³ I claim that causal explanations involve a notion of causal necessity. As Kit Fine writes,

Suppose that one billiard-ball hits another. We are then inclined to think that it is no mere accident that the second billiard-ball moves. Given certain antecedent conditions and given the movement of the first ball, the second ball *must* move. And the ‘must’ here is the *must* of natural necessity.¹⁴

What Fine calls *natural* necessity, I call *causal* necessity. This causal necessity seems to be a necessary condition of causal explanation. That is to

¹²Fine 2002, p. 254.

¹³Salmon 1984.

¹⁴Fine 2002, p. 256.

say:

(1.2.1) p causally explains q only if p and q and it is causally necessary that if p then q .

In 1.2.1 I take there to be a causal explanatory connection between the propositions that p and that q . Nevertheless, 1.2.1 is neutral on the question of the *fundamental* nature of causation. In particular, it does not imply that causation fundamentally connects facts, as defended by Mellor¹⁵ rather than relating events, as defended by Davidson.¹⁶ It is compatible with 1.2.1 that causal necessity is to be analysed in terms of a causal relation between events.

1.2.2 Inferential Explanations

Consider the claim that:

(1.2.2) Socrates is mortal because Socrates is a man and all men are mortal.

1.2.2 is very naturally taken to express an explanation in the *inferential* mode. Crudely put, the claim that there is an inferential mode of explanation corresponds to the view that some explanations are *arguments*.¹⁷ The obvious position to take here is that inferential explanations involve logical necessity, that is, entailment. So I claim:

¹⁵Mellor 1991.

¹⁶Davidson 1967.

¹⁷Cf. Salmon 1984, p. ix.

(1.2.3) p inferentially explains q only if p and q , and p entails q .

Though it is often tempting to deny that a certain argument's premisses provide an explanation of its conclusion, we can account for those intuitions in a less radical way by claiming instead that, in many contexts, it is not inferential explanation but some other mode of explanation that interests us. In science, we are often interested in causal explanations. But in logic and mathematics, an appropriate explanation may come in the form of a proof consisting of a deductively valid argument. So we ought to accept that there is an inferential mode of explanation.¹⁸

1.2.3 Evidential Explanations

It is well recognized that there are *epistemic* or *evidential* uses of 'because', such as:

(1.2.4) The volcano is about to erupt because it is smoking.

(1.2.5) She is at home because her light is on.

One way to understand these claims is to accept that they are good explanations, but to insist that they do not wear their explananda on their sleeve. For example, we can take 1.2.4–1.2.5 to be elliptical expressions of something closer to the following:

(1.2.6) It is epistemically justifiable to believe that the volcano is about to erupt because it is smoking.

¹⁸If there is an inferential mode of explanation, it may be that p can be an inferential explanation of itself, but *only if* p is true.

(1.2.7) It is epistemically justifiable to believe that she is at home because her light is on.

Here I have given impersonal expression to the justifiability of a belief, but we could also index the justifiability both to a person and to a given body of knowledge or evidence. Indeed, on one development of this view, we would need to claim that 1.2.4–1.2.5 wear neither their explananda nor their explanantia on their sleeve, since both are elliptical for something closer to the form of:

(1.2.8) It is epistemically justifiable *for S* to believe that the volcano is about to erupt because *S* knows (or at least, believes with epistemic justification) *inter alia* that the volcano is smoking.

On this view, epistemic or evidential explanations are elliptical for a certain kind of explanation given in terms of the epistemic justification of a belief.¹⁹ If that is correct, then it is natural to think that these explanations require epistemic necessity as follows:

(1.2.9) *p* *epistemically* explains *q* only if *p* and *q* and it is epistemically necessary that if *p* then *q*.

Alternatively, we might account for these explanations in terms of evidence, as follows:

(1.2.10) *p* evidentially explains *q* iff *p* and *q* and *p* is evidence that *q*.

¹⁹For the sake of a clear distinction, I am taking epistemic justification to be non-normative. If epistemic justification is normative, that shows only that normative explanation is more widespread and more interesting.

1.2.9 associates epistemic explanation with *epistemic necessity*, the necessity of what *must* be true given one's *knowledge*. 1.2.10 associates evidential explanation with *evidence*. It is, I concede, not obvious whether *p*'s being evidence for *q* can be understood as a mode of necessity. Suppose, however, that we take a probabilistic view of evidence, perhaps relative to an agent's knowledge.²⁰ One plausible constraint to place on an account of evidential explanation is that, if *p* evidentially explains *q*, then *p* must not simply raise the probability that *q* (say from one in a billion to one in a million); rather, it must make *q* more probable than improbable, and hence raise the probability of *q* above a half. And, at a stretch, one might say that *p*'s making it probable that *q* provides a weak, but recognisable, sense of necessity.

1.2.4 Conceptual and Metaphysical Explanations

Consider the following explanations:

(1.2.11) John is a bachelor because he is an unmarried man.

(1.2.12) Figure *A* is a triangle because it is a closed figure of three straight sides.

What mode of explanation might underpin these claims? One view would be that these explanations are primarily conceptual or semantic in nature. I do not profess to have a full account of the workings of these conceptual explanations, but a starting view would be as follows:

²⁰Cf. Williamson 2000, Ch. 9.

(1.2.13) p conceptually explains q only if p and q , and it is conceptually necessary that if p then q .

Of course, the tenability of this view depends on the tenability of a view of conceptually necessary (or, in other terminology, *analytic*) truths. It is no part of my task to defend that view. I note only that such views have received both significant defences and significant attacks.²¹

An alternative account of 1.2.11–1.2.12 would re-conceive what we originally took to be conceptual explanations as, in fact, ontological or metaphysical explanations. Writing on the topic of analyticity, Williamson writes:

The sentence ‘Vixens are female foxes’ is in no useful sense about the word ‘vixen’ or any other words; it is about vixens, if anything.²²

Similar things may be said about 1.2.11–1.2.12. Rather than taking these claims to be about concepts, we can take them to be about the metaphysics of what is to be an unmarried man and what it is to be a triangle. Adopting the metaphysical reading of these explanations, we might say that John is a bachelor because he is an unmarried man, and being an unmarried just is *what it is* to be a bachelor.

The metaphysical readings of 1.2.11–1.2.12 put them much closer to these claims:

(1.2.14) Kind A is water because kind A is H_2O .

²¹Critics of conceptual truths include Quine 1951; Williamson 2007, Chs. 3–4. Defenders include Jackson 1998.

²²Williamson 2007, p. 49.

(1.2.15) The Star of Africa is a diamond because it is an isometric-hexoctahedral crystal lattice of carbon atoms.

What mode of explanation might underpin these explanations, now that they are understood metaphysically? I suspect that a plurality of metaphysical connections that might be involved, but that what they have in common is metaphysical necessity. In general, then, my account of metaphysical explanation reads:

(1.2.16) p metaphysically explains q only if p and q , and it is metaphysically necessary that if p then q .

Various metaphysical connections might underpin a metaphysically necessary connection. For example, the metaphysical necessity that connects kind A 's being water with kind A 's being H_2O is, presumably, the identity of the kinds *water* and *H₂O*. In other cases, the metaphysical necessity might be underwritten in other ways. For example, the metaphysical necessity that connects Figure A 's being a closed figure of three straight sides with Figure A 's being a triangle may be the fact that triangles are *constituted by* three straight sides arranged in a closed figure. There are several possibilities here for the defender of metaphysical explanation.

1.2.5 Normative Explanation

I defended above the legitimacy of a number of plausible non-normative modes of explanation: causal, inferential, evidential, conceptual and/or metaphysical. I take it that the *superficial* distinctness of at least some of

these modes of explanation is self-evident. If an argument is needed for this claim, let it be that there is such a thing as causal explanation, but there is also such a thing as non-causal explanation. For example, the claim that Donald quacks because Donald is a duck and all ducks quack is a genuine explanation in some mode, but it is not a causal explanation. QED for superficial explanatory pluralism.

I turn at last to the idea that there is a distinctively *normative* mode of explanation. One problem to be faced in characterizing normative explanation is that, just as normative concepts often seem indefinable²³, so, too, does normative explanation. How, then, to characterize it? In § 1.3, I will characterize normative explanation by reference to its theoretical rôle, in particular, its rôle in moral theory and in the conception of the good-willed or virtuous agent. But, it is better begin with some illustrative examples:

(1.2.17) That Sarah promised to help explains why she ought to help.

(1.2.18) Sue's revealing that information was wrong because it was unnecessarily cruel.

(1.2.19) That Bob's taking the job is for the best makes it the case that Bob ought to take the job.

It may be possible to read some of these claims as involving non-normative modes of explanation, but I intend them all normatively. And so understood, I take their intended normative character to be self-evident.

1.2.17–1.2.19 are examples of explanations of why an act has some *over-*

²³Cf. Moore 1993, Ch. 1.

all or *all things considered* moral or normative status. There are also normative explanations for facts of non-overall moral or normative status. For example, we could claim that Sue's revealing that information was *pro tanto* or *prima facie* wrong because it was unnecessarily cruel. This raises the question of how far the phenomenon of normative explanation extends. Does it, for example, encompass evaluative explanations, as in these examples:

(1.2.20) Bob's phoning an ambulance was good because Sue was in severe pain.

(1.2.21) The murderer's actions were bad because they caused loss of life.

(1.2.22) The holiday was good because it was a lot of fun.

And might it also extend to explanations of *areteic* or *virtue-related* facts?

E.g.:

(1.2.23) John's donating to charity was benevolent because it saved a child's sight.

(1.2.24) Bob's phoning an ambulance was kind because Sue was in severe pain.

(1.2.25) Sarah's keeping John's secret was faithful because she put a high value on their friendship.

These explanations of evaluative and areteic facts seem to me to be of a

more variable normative character than the explanations of normative facts. Furthermore, it would be somewhat surprising if a single mode of explanation was straightforwardly involved in explaining these varied phenomena. For these reasons, I will henceforth assume that normative explanation does not encompass explanations of evaluative and areteic facts.

The obvious question to ask, once the normative mode of explanation is recognized, is how it relates to the modes of explanation already discussed. It is evident that normative explanation is *superficially* distinct from each of them. None of my examples of non-normative modes of explanation had a normative explanandum of any kind, but they were genuine explanations. So, normative explanation is superficially distinct from each of them. But, this is an easy victory. A more interesting claim to defend is that the normative explanation mode of explanation is distinct from the other modes we have encountered, even when those other modes take a normative explanandum. Is there really more than one way to explain a normative fact?

Consider, to begin, whether normative explanation might be reducible to causal explanation. This seems false for several reasons. First, it is doubtful whether causal explanations of normative facts are even possible. Consider what might be the nearest thing. Suppose that *A* is motivated to kill person *B* by the desire wrongly to kill him. If desires can causally explain actions, is it perhaps that *A*'s desire causally explains why *A* wrongly kills *B*? I claim that it does not. That *A* has a desire wrongly to kill *B* may causally explain why *A* kills *B*. But *A*'s desire does not necessarily explain why *A*'s killing *B* is wrong. Second, it seems possible that the

rightness or wrongness of a given act might be normatively explained by its future consequences. So, unless there can be causation backwards through time, normative explanation is not reducible to causal explanation. Third, it may be that causal explanations essentially involve a spatio-temporally continuous causal process, so that, strictly speaking, there is no such thing as ‘causation at a distance’.²⁴ However, in the moral case, the fact that last week I promised to help you mow your lawn may explain why I ought now to help you mow your lawn, without any intervening process at all. Causation may necessarily be proximal, but normative explanation can be at a distance. Fourth, though Donald Davidson has argued that reasons are to be understood as causes, this project refers to *motivating* rather than *normative* reasons.²⁵ So, the Davidsonian project does not support the idea that normative explanation may be reduced to causal explanation. I conclude, consequently, that normative explanation is both superficially and deeply distinct from causal explanation.

Consider next the following explanation:

(1.2.26) Oswald’s killing of JFK is wrong because

1. either Oswald’s killing of JFK is wrong or chewing gum is not sticky; and
2. chewing gum is sticky.

Here the explanantia *inferentially* explain the explanandum, but they do not *normatively* explain it. Hence normative and inferential explanations

²⁴Cf. Salmon 1984.

²⁵Motivating and normative reasons are discussed further in Chapter A.

are superficially distinct.

Consider now some epistemic or evidential examples:

(1.2.27) Jim ought to give the money back because if he does not it will prey on his conscience.

(1.2.28) It would be wrong to set that cat on fire because everyone would be horrified.

(1.2.29) Sarah ought to take the job because a wise guru told her to take it.

I take 1.2.27–1.2.29 to be epistemic explanations, but they are in all cases poor normative explanations, if normative explanations at all. The fact that everyone would be horrified may be good evidence that it would be wrong to set a cat on fire, but the normative explanation surely lies elsewhere, having to do with avoiding unnecessary agony to a conscious creature. Hence the normative and evidential modes of explanation are superficially distinct.

Finally, consider these claims:

(1.2.30) Oswald's killing JFK was wrong because he ought not to have done it.

(1.2.31) Jim ought to attend the meeting because that is what there is most reason to do.

I claim that 1.2.30–1.2.31 can be understood as offering either conceptual or metaphysical explanations. 1.2.30 seems to depend on a conceptual connec-

tion between *wrongness* and *ought not*. Similarly, 1.2.31 might be claimed merely to rest on an analysis of either the concept (or the property) *ought* into the concept (or the property) of *most reason*.²⁶ Might they also be normative explanation? I claim that they cannot, but my argument is given in § 2.2.2. In Chapter 4, I argue more generally against the reduction of normative explanation to conceptual necessity.

1.3 The Nature of Normative Explanation

Let us review the results of previous sections. I have argued that explanations are the values of explanatory functions. I have also defended a weak version of explanatory pluralism, according to which there is a plurality of explanatory functions that are superficially distinct. Finally, I have argued that there is distinctive normative mode of explanation that is at least superficially distinct from any other mode of explanation. It is this normative mode of explanation that now forms the central focus of my thesis. In this section, I aim to do three things. First, I aim to define in precise terms the question to which this thesis is addressed. Second, I aim to characterize more precisely the normative mode of explanation. In so doing, I hope, third, to show why normative explanation merits investigation.

I introduced normative explanation with some examples of explanations of *overall* or *all things considered* normative or moral facts, that is, explanations of why an act is overall right or wrong. I also said that there might be normative explanations of facts of *non-overall* normative or moral status

²⁶This conceptual analysis of *ought* is defended in e.g. Parfit 2007, p. 21.

i.e. explanations of why an act is *prima facie* or *pro tanto* right or wrong. Though these two aspects of normative explanation are clearly related, it is by no means clear that they deserve the same treatment. For this reason, I propose to focus on normative explanations of overall normative facts. Henceforth, then, by *normative explanations* I will mean explanations, in the normative mode, of facts of overall normative status. I will now define these facts more precisely.

I start with the idea that there are a number of categories of overall *moral* or *normative* status. These categories are typical of what are sometimes called *verdictive* or *all things considered* normative judgements. According to a standard view, there are three such categories: *obligatory*, *permitted*, and *wrong*. However, we can conceive of these categories as operating in different ways. On one conception, the overall normative categories correspond to monadic properties of an act. On a second conception, the deontic categories are taken to be relations between an agent, an act-type, and perhaps a time.²⁷ On yet a third conception, the deontic categories are taken to be propositional operators. I aim as far as possible to say neutral on these different conceptions, but for consistency of presentation, my official policy will be to take the overall normative categories to correspond to properties of actions. I will therefore take normative explanation to be explanations of why an act has a certain overall normative property.

Next, I need to clarify what I mean by an *act*. I am interested in explanations of the normative status of act-tokens rather than act-types. That is to say, I am interested in why a particular instance of, e.g. a lying

²⁷See e.g. Geach 1991.

act, is wrong, rather than in explanations of why a *lie*, as an act-type, is wrong. This working stipulation is not incompatible with the view, held by some, that explanations of right and wrong fundamentally attach to act-types rather than tokens. If that view is correct, we should expect an account of the explanation of right and wrong in token acts to make that evident.

The notion of a token act is itself in need of clarification. The word ‘act’ exhibits a common form of act/object ambiguity.²⁸ It is well recognized that the word ‘belief’ can refer either to the particular mental act or state of *believing*, or to the object of that believing, the *thing believed*. So, too, the word ‘act’ can refer either to the mental state or activity of an agent, namely, the agent’s *doing* of something, or to the object of that mental activity, the *thing done*. The applicability of the act/object distinction for acts is too rarely recognized in ethics for there to be a standard view of which is the primary focus of moral assessment. Indeed, I think it is both possible and desirable to attribute overall normative status to *acts* in both senses. For this reason, I will assume that a normative explanation always applies to an *act* in both senses. That being so, there will be no need to mark this distinction.

Combining these clarifications, I can now define an overall normative fact as the proposition expressed by a sentence of the schematic form ‘*Da*’, where ‘*a*’ names a token act (whether a doing or a thing done), and ‘*D*’ is any predicate expressing a property of overall normative status. Normative explanation can now be defined as explanation, in the normative mode, of

²⁸This claim is made by e.g. Hornsby 1997.

overall normative facts. This leads me to formulate the central question to be addressed in this thesis as follows:

My Central Question What is the nature of normative explanation?

My central question encompasses many others. What are the truth conditions of normative explanations? Is normative explanation reducible to another mode of explanation? What formal properties does it have? I return to these questions in the next chapter. My task in the remainder of this chapter will be to clarify further my central question and the normative mode of explanation it concerns.

My central question asks after the nature of normative explanation. I contrast questions about the *nature* of x with questions about our *concept*(s) of x . While an investigation into the nature of some entity presupposes and makes use of our conceptual contact with it, I take there to be an obvious sense in which the entity itself is prior to any concept of it. For example, when a concept of x is incomplete, confused, vague, or otherwise defective, it may be improved by further investigation into the nature of x . In the Anscombian metaphor, there is concept-to-entity *direction of fit*. For this reason, my loyalties lie with an investigation into the nature of normative explanation. I hope to *improve*, and not merely *describe* our concept of it.

The distinctive focus of my central question becomes clearer if we compare it to some related questions asked by two other writers, Ross and Broad. In *The right and the good*, Ross writes:

The real point at issue between hedonism and utilitarianism on the one hand and their opponents on the other is not whether

‘right’ means ‘productive of so and so’; for it cannot with any plausibility be maintained that it does. The point at issue is that to which we now pass, viz. whether there is any general character which makes right acts right, and if so, what it is.²⁹

Ross’ question is echoed by Broad:

Let us call those non-moral characteristics whose presence in anything confers rightness or wrongness on it *right-making* and *wrong-making* characteristics. [...] Now an extremely important question is whether we can discover any kind of systematic unity among all these various right-making and wrong-making characteristics. Can we reduce them to a few fundamental ones? Can we perhaps reduce them all to a single fundamental one?³⁰

On a natural interpretation, Ross and Broad seek to know something about the facts which normatively explain why token acts are right and wrong. They seek to know, for example, which those facts are, and whether there is any ‘systematic unity’ among them. In short, their questions concern the *explanantia* of normative explanations. By contrast, the focus of my question is with the normative mode of explanation itself, and not, or at least not only, with its explanantia. So my question is a different one.

Here is another way to understand the difference between Ross’ and Broad’s questions and my own. Ross and Broad are concerned with discovering which first order normative theory is correct. But, as I argue below, my aim is instead to understand the normative mode of explanation which I take all first order normative theories to share in common. In consequence, I aim as far as possible to avoid taking any stand on the question of which facts normatively explain why acts are right or wrong — to avoid, that is,

²⁹Ross 2002, p. 16.

³⁰Broad 1946, p. 103.

taking a stand on the first order questions asked by Ross and Broad. In discussing the nature of normative explanation, I will instead deploy a variety of examples, drawn from different first order theories. My strategy can certainly be criticized for taking examples from a variety of theories which are incompatible among themselves, but it has the virtue of spreading risk and helping to ensure that examples are not drawn from too narrow a range of possibilities.

Having clarified my central question, I return to the topic of normative explanation and its importance for ethical theory. In § 1.2.5, I introduced normative explanation using some examples. While examples are useful, it will help my investigation to characterize normative explanation as fully as possible. In the remainder of this chapter, I argue that normative explanation can be characterized through its central rôle in normative theory, and its connection to the motivation of virtuous agents. Understood in this way, normative explanation is of central importance to ethical theory.

1.3.1 Normative Explanation and Normative Theories

Jonathan Dancy has recently written:

Some normative facts, then, are more complex than the simple fact that one ought to do this; they contain that fact, but they also contain what makes that simple fact the case. Such facts are of this form: that p makes it the case that one ought (or has reason) to act in way w . These metafacts are facts about some matter of fact and about its making a difference to how to act. They constitute direct answers not only to the practical question

what to do, but also to the question why. It is these metafacts that I think of as the central normative facts, by reference to which the normativity of all others is to be understood.³¹

In this section, I build on Dancy's idea that *explanatory* facts have a central rôle in normative theory.

One way of characterizing normative explanation is by reference to its rôle in first order normative ethical theory. This is the level of theory that corresponds to what is sometimes called a 'theory of the right'. Of course, by a 'theory of the right' what is meant is a theory of all the overall normative categories. However, first order normative theories can be understood in different ways, so let me make some distinctions.

A theory is a set of propositions. *Moral* or *normative* theories come in various forms. (I remind the reader that I use 'moral' and 'normative' interchangeably.) While I believe these different types of theory can ultimately be differentiated according to their content, we can helpfully distinguish them in a more intuitive way by reference to the questions they aim to answer. So let me distinguish some of the important moral questions.

One important question in moral theory concerns the nature of the overall normative properties. Accordingly, one type of moral theory is a theory of the nature of these properties. In *What we owe to each other* Scanlon describes his contractualism as an account of the property of wrongness.³² Similarly, Moore in *Principia Ethica* maintained that the property of being right and the property of producing the most good were

³¹Dancy 2006*b*, p. 137.

³²Scanlon 1998, p. 12.

identical.³³ Both Moore and Scanlon, then, offer us theories of the nature of the normative properties.

A second important question in moral theory is that of which acts are right, wrong, and permissible, or more generally, which acts belong to which overall normative categories. We can understand this question as asking after the extensions of the various overall normative predicates, ‘*x* is obligatory’, ‘*x* is wrong’, ‘*x* is permissible’. A second kind of moral theory, then, is a theory of the extensions of the overall normative predicates. At its most ambitious, a theory of this sort might aim to specify their extensions in all possible worlds.

According to many, however, one of the central questions in moral theory is, quite simply, that of what *makes* acts right, that is to say, what *explains* why they are right.³⁴ Mark Timmons expresses this view:

The main theoretical aim of moral theory is to discover those underlying features of actions, persons, and other items of moral evaluation that make them right or wrong, good or bad.³⁵

Timmons view of moral theory is broader than my own at this point because it also includes the evaluative question of what makes things good or bad. But it is enough for present purposes that Timmons and I are in broad agreement in the domain of normative theory. One kind of moral theory,

³³Moore 1993, § 17, p. 17.

³⁴The idea that the question of what makes acts right and wrong is an important one has not gone unchallenged. In particular, it will be noted that this question is formulated using moral concepts. See Crisp 2006, Ch. 1, esp. 16–17 for a defence that the important questions should be asked using only the concept of a normative reason. I discuss normative reasons in Chapter 5.

³⁵Timmons 2002, p. 4.

then, is a theory of what normatively explains why acts belong to the various deontic categories. These are the *first order normative theories*.

A first order normative theory is a theory of what explains, in the relevant normative mode, why acts belong to the various overall normative categories. What would such a theory precisely look like? One account might be this: a first order normative theory is a set of propositions which state explicitly what normatively explains why each token act, in each possible world belongs, to the overall normative categories that it does belong to, and nothing more. So understood, a normative theory must rest upon some specification of all and only the genuine normative categories. Combining these requirements we arrive at the following definition of a normative theory. A normative theory is a set of propositions comprising:

1. A proposition which states explicitly that some set with specified members is the set of all and only the overall normative categories.
2. A number of propositions which state explicitly what normatively explains why each possible act, in each possible world, belongs to whichever overall normative categories it does, in fact, belong to in that possible world.
3. Nothing else besides.

The ‘Nothing else besides’ condition serves to rule out irrelevant information.³⁶ I will illustrate my account of first order normative theories with an example. Let ‘optimal’ mean ‘not bettered by any relevant alternative’, and

³⁶My ‘nothing else besides’ condition is identical in function to the ‘stop’ clause described in Jackson 1998, pp. 12–13.

let ‘sub-optimal’ mean ‘bettered by at least one relevant alternative’. One might define a simple version of maximizing direct act consequentialism as follows:

(1.3.1) **Maximizing Direct Act Consequentialism (MC)**

1. The overall normative categories are: obligatory, permissible, wrong.
2. Necessarily for all acts x :
 - a) x is obligatory iff that x is uniquely optimal normatively explains why x is obligatory.³⁷
 - b) x is permissible iff that x is optimal normatively explains that x is permissible.
 - c) x is wrong iff that x is sub-optimal normatively explains that x is wrong.

The claim that a normative theory is to be defined in the manner I have described is significant for three reasons. First, it provides a practical model of how a normative theory ought to be defined. Second, it entails that normative explanation is essential to the identity conditions of normative theories. Normative theories are sets of propositions, and set identity is determined by membership. So, it follows from my account of definition that a normative theory x and a normative theory y are identical iff:

³⁷Some might prefer to state what makes acts obligatory/permissible/wrong as a necessary condition only e.g. x is obligatory only if that x is uniquely optimal normatively explains why x is obligatory. However, I argue in § 2.2.1 that normative explanation is factive, so the ‘only if’ claim entails the ‘iff’ claim.

1. x and y propose identical sets of the genuine normative categories
2. x and y propose identical sets of propositions about what normatively explains why belong to each of the genuine normative categories
3. x and y include nothing more than is required by 1 and 2.

The view that normative explanation is essential to the identity conditions of normative theories can be given further support. To avoid prejudging the issue, call a normative theory as defined above an ‘explanatory theory’. It is widely accepted, on the basis of Plato’s arguments in the *Euthyphro*³⁸, that distinctness of explanatory theory entails distinctness of normative theory. That being so, all that remains to be shown is that identity of explanatory theory implies identity of normative theory. But it is hard to see how this claim could be false: what else could the identity of a moral theory depend on? If two scientific theories acknowledged the existence of the same entities and offered identical explanations of observed phenomena, it would be hard to deny that they were identical. The same goes for normative theories.

The third reason why this account of moral theories is important is that it entails that *explanatory over-determination* is not a harmless phenomenon. Explanatory over-determination occurs when there is more than one normative explanation of a particular normative fact. This kind of over-determination is sometimes thought to be harmless but if, as I have argued, normative theories are sets of propositions concerning normative explanations, then adding a normative explanation to an existing theory or

³⁸We can ask the Euthyphro question using the concept of normative explanation, whether or not that was Plato’s own intended reading. See 1.3.2 for further discussion.

subtracting one from it results in a *distinct* normative theory. One ought, for this reason, to be wary of permitting over-determination without good reason.

1.3.2 Higher Order Normative Theories

In § 1.3.1 I discussed the rôle of normative explanation within first order normative theory. As an adjunct to that discussion, I want briefly to clarify a frequently overlooked distinction between first order normative explanations and higher order normative explanations.

A first order normative explanation is an explanation of the form ‘ p normatively explains Da ’ where a is an act and D is an overall normative property. First order normative explanations answer both the question of what ought (or ought not) to be done, and the question of why it ought (or ought not) to be done, where this why-question is understood as a request for normative explanation. For example: One ought to do act a , and one ought to do it (normatively) because p . But there is a further question one might raise here: why does p normatively explain why one ought to do act a ? Here we are asking for a second order explanation of a first order explanation. Lance and Little draw this distinction as follows:

We can begin by distinguishing two different tasks that purported principles have been asked to play in morality. Broadly put, *normative principles* purport to articulate which considerations count as good- or bad-making, right- or wrong-making. In contrast to ontological claims about what, as it were, *make* good-making features good-making — a divine commandment, a Platonic Reality, the output of some idealized contract — nor-

mative principles aim to set forth those that do so count.³⁹

Where Lance and Little draw a distinction between *normative* and *ontological* principles, I draw a distinction between first and second order normative explanations.

By a second or higher order normative explanation I mean an explanation, which *may or may not* be in the normative mode, of a first order normative explanation. We can make explicit the contrast between first and second order normative explanations as follows:

First Order Schema That *p* normatively explains that *Da*.

Second Order Schema That *q* explains that (that *p* normatively explains that *Da*).

The brackets in the Second Order Schema clarify the scope of the higher order explanation.

Second order normative explanations are important for several reasons. As I understand it, it is a second order question that is addressed by Prichard in his classic paper, “Does moral philosophy rest on a mistake?” Prichard writes:

Any one who, stimulated by education, has come to feel the force of the various obligations in life, at some time or other comes to feel the irksomeness of carrying them out, and to recognize the sacrifice of interest involved; and, if thoughtful, he inevitably puts to himself the question: ‘Is there really a reason why I should act in the ways in which hitherto I have thought I ought to act?’⁴⁰

³⁹Cf. Lance and Little 2006*b*, p. 569.

⁴⁰Prichard 2002, p. 7.

Prichard's question can be understood in several ways. On the simplest reading, Prichard is simply questioning his moral beliefs, asking a general question of the form 'I believe I ought to ϕ , but ought I to ϕ '? This reading lacks interest.

A second interpretation of Prichard's question takes him to recognize the *moral* status of certain acts as obligatory and wrong, but to ask whether one is *normatively required* to perform the obligatory acts and refrain from the wrong acts. Understood in this way, Prichard is asking whether moral requirements are normative requirements. However, there are those, myself included, who take moral requirements to be a subset of normative requirements, and obviously so. For us, Prichard's question asks, entirely trivially: are the normative requirements of morality normative requirements? For this reason, I find this second reading of Prichard's question no more interesting than the first.

The better reading, I suggest, is that Prichard is asking a second order question. 'Stimulated by education', one comes to recognize that there are certain first order normative facts of the form 'I ought/not to ϕ because p .' And though one recognizes the normative character of these first order facts, one also recognizes the irksomeness of being under those first order obligations. This leads one to question, not why one ought (or ought not) to ϕ , for this answer is already given: one ought to ϕ (normatively) because p ; but rather, one is led to ask why the fact that p normatively explains why one ought (or ought not) to ϕ ? One seeks a second order explanation of a first order normative explanation.

Before Prichard's question can be answered, we must ask how second

order normative questions are to be understood. In particular, what mode of explanation is being employed at the second order level? One possible answer is that it is the normative mode of explanation that is reappearing at the upper level. This appears to be Prichard's own position. According to Prichard, the only possible answers to the question of why one ought to act as morality requires are either (1) because it is to one's own advantage; or (2) because doing so some realizes some good. Prichard objects that adopting the first alternative is to 'presuppose an intermediate link, viz. the further thesis that what is good ought to be.'⁴¹ However, adopting the second alternative involves the fallacy that to base the rightness of an act upon its intrinsic goodness implies that the goodness in question is that of the motive', whereas 'in reality the rightness or wrongness of an act has nothing to do with any question of motives at all.'⁴² Prichard concludes from this argument that moral philosophy rests on a mistake: if morality is not an illusion, some answer is needed to the second order normative question; but all answers to this question are false; hence morality *is* an illusion.

I reject what I take to be Prichard's assumption that his second order question concerns the normative mode of explanation. I reject it because it is doubtful that normative explanations can themselves be normatively explained. Normative explanations seem capable of explaining only why something ought to be the case. So a second order normative explanation of a first order normative explanation would seem capable of explaining only

⁴¹Prichard 1949.

⁴²Prichard 1949.

why there *ought* to be first order normative connections. But that will not answer Prichard's question. The question 'Is there a reason why I ought to do what I ought to do' cannot be answered by claiming that *there ought to be* a reason.

Prichard's mistake is to confuse two senses of the question 'Is there a reason why *p*?' Understood in one way, this question asks for a normative explanation why *p*. But as John Broome points out,

Another confusing feature of English is that the explanation of why a fact obtains is also called 'the reason' why it obtains. Here, 'the reason' is used in a non-normative sense. In this sense, it may be applied to any explanation, whether of a normative or a non-normative fact. The reason why pigs cannot fly is that they have no wings.⁴³

We can avoid Prichard's own mistake if we take second order normative explanations to involve a *non-normative* mode of explanation at the upper level. That is, what is needed is a non-normative explanation of why certain facts normatively explain why we ought to act in certain ways. To provide further support for this improved reading of Prichard's question, let me show how a number of familiar theories can be understood as offering higher order, *non-normative* explanations of normative connections.

I claim that many moral theories can be understood in two ways. They can be understood as first order normative theories, offering normative explanations of why acts are right and wrong. And they can be understood as second order normative theories, offering a non-normative explanation of first order normativity. Consider for a first example the divine command

⁴³Broome 2004a, p. 34.

theory. Taken as a first order theory, divine command theory claims, *inter alia*, that:

1st Order Divine Command Necessarily for all acts x : x is obligatory
iff the fact that God's commands x normatively explains why x is
obligatory.

Thus, the starving should be fed (when they should) because God so commands, promises should be kept (when they should) because God so commands, and so on. One obvious objection to this theory is that it appears to deny any normative explanatory force to the bare facts that people are starving, that promises were made, and so on. I shan't linger on these objections; I mention them merely in order to show what improvements await if we adopt a second order divine command theory.

Suppose we accept that the fact that someone is starving normatively explains why one ought to feed them, and that the fact that one has made a certain promise to ϕ normatively explains why one ought to ϕ . Why, one might ask, do these normative explanations hold true? One schematic answer to this question would be this: in every case where one ought to ϕ because p , this is so because God has commanded, on the basis that p , that we ϕ , and God's so commanding is constitutive of first order normative explanation.⁴⁴ That is to say, God's commanding, on the basis that p , that one ϕ , is *metaphysically constitutive of*, or perhaps even *identical with*, the fact that that q normatively explains why one ought to ϕ . In short, norma-

⁴⁴I use the phrase 'God's commanding, on the basis that p , that one ϕ ' to make clear that p is God's own motivation for so commanding. It does not follow that we are commanded to ϕ with that same, or any other, motivation.

tive connections are identical to, or constituted by, divine command.⁴⁵

Here is a second example. It is clear that the moral position articulated in Plato's *Republic* is a form of egoism, and many have thought this was problematic. The problems, I suggest, derive from mislocating Plato's egoism at the first order level. Taken as a first order theory, egoism claims, *inter alia*, that:

1st Order Egoism Necessarily for all agents a and acts x : a ought to do x iff the fact that doing x is in a 's best interest normatively explains why a ought to do x .

The problem with reading first order egoism into the *Republic* is that Plato's egoism is intended as a foundation for justice, and the just person is someone who, at least on occasion, is motivated to help others for their sake rather than for his own. Now, this is a consistent view, and some have tried to defend it by claiming that it is in a just person's interest to be motivated to help others for their sake, rather than for his own.⁴⁶ However, as a reading of Plato this has a fatal flaw. The just person is one who acts from knowledge of the 'Form of the Good'. Whatever else this means, I think it must mean that the just person helps others for their own sake because he recognizes that their interest, rather than his own, normatively explains why he ought to help. Hence Plato cannot be a first order egoist.

The better hypothesis, I suggest, is that Plato's egoism is an egoism of

⁴⁵Adams 1979 defends a divine command theory according to which wrongness is identical with the property of being contrary to the commands of a loving God. My version of divine command ethics is significantly different in taking God's commands to have a structure that is constitutive of first order normative explanation.

⁴⁶Cf. Annas 2008.

the second order level. Suppose I accept that on occasion I ought to help others (normatively) because it is in their interest. Why do facts about what is of interest to other normatively explain why I ought to help them? A schematic answer would be this: in every case where one ought to ϕ in light of its being the case that p , this is so because ϕ -ing in light of its being the case that p is in one's interest, and facts about whether doing something in a certain light is in one's interest are metaphysically constitutive of, or identical with, first order normative connections. In short, normative connections are identical to, or constituted by, considerations of self-interest. Understood in this way, Plato's second order egoism is compatible with genuine, first order altruism.

Consider, as a final example, the version of a Humean theory of practical reason recently defended by Mark Schroeder:

Reason 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 .⁴⁷

Schroeder comments: 'Reason tells us *what it is* for R to be a reason for X to do A . . . it gives us a *constitutive* account, or an analysis, of what it is for R to be a reason for X to do A . . .'. And he continues:

[...] desires have to serve in the explanation of *every* reason, because desires are part of the correct *analysis* of reasons. This analysis is a *reductive* one. It analyzes *reasons*, a normative category, wholly in non-normative terms.⁴⁸

⁴⁷Schroeder 2007, p. 59.

⁴⁸Schroeder 2007, p. 60.

Recast in my own terms, what Schroeder is offering is a second order, non-normative explanation of first order normative connections.

No doubt these second order normative explanations require further elaboration and defence, but I think their interest is already clear. I leave this discussion of second order normative theories with some final observations.

My first observation concerns *ethical particularism*, or rather, since particularism is a family of theories, it concerns one issue often discussed under that name. In § 1.3 I quoted Ross' question of 'whether there is any general character which makes right acts right, and if so, what it is', and Broad's question of 'whether we can discover any kind of systematic unity among all these various right-making and wrong-making characteristics'. Moral particularism is often associated with the claim that there is no, or very little, of this 'systematic unity', but this view has both its defenders and critics.⁴⁹ Against this claim, Jackson, Pettit, and Smith have argued:

The only plausible explanation of our capacity to use evaluative predicates to mark distinctions in the ways things are is the existence of a descriptive pattern unifying all cases of right action.⁵⁰

One overlooked position, so it seems to me, is that there is a systemic unity among the various right-making and wrong-making characteristics, but that the relevant pattern occurs only at the level of second (or higher) order explanation. Suppose we accept for example, that we have a long list of first order obligations to perform token acts each with their particular normative

⁴⁹See Hooker and Little 2000 for a useful survey.

⁵⁰Jackson, Pettit, and Smith 2000, p. 99.

explanations: we ought to feed this starving person because he is hungry, we ought to help this other person for his own sake, we ought to take a particular holiday because it will be fun, and so on. One can deny that there is any systematic unity here at the first order level, and still be committed to systematicity at the second level. I have already described three such theories that have the resources to do exactly this: divine command, Platonic egoism, and Schroeder's Humeanism. What unifies the various first order facts is that e.g. each of them is representative of a divine command; or that each of them is an instance of the exercise of human virtue in which our welfare consists; or that each of them is representative of a desire that we have for some object based on one or more of its qualities.

My final point regarding the importance of higher order normative explanations is that explanations are *factive*⁵¹ So if we have a non-normative second order explanation of a first order normative explanation of why I ought to ϕ , it follows that I ought to ϕ . Thus, second order normative explanations can be of direct first order significance.

I have clarified the distinction between first and higher order normative theories the better to delineate the first order theories that I take to characterize normative explanation. I proceed now to my secondary characterization of normative explanation in terms of virtuous agency.

1.3.3 Normative Explanation and Virtuous Agents

Normative explanation has interesting connections with virtuous agency. At the shallower end of these connections is the idea that a virtuous agent

⁵¹This follows from the truth table for explanatory connectives given in 1.2.

will be concerned to know what, in particular cases at least, normatively explains why token acts are obligatory, wrong, and permissible, the better to know which token acts *are* obligatory, wrong and permissible. On this view, our knowledge of moral facts is often based on our knowledge of the presence or absence of features which normatively explain them.

A second connection between normative explanation and virtuous agency arises in the use to which a virtuous agent puts his moral knowledge. Once a virtuous agent knows which features normatively explain why certain acts are right or wrong, he can employ this knowledge in his practical deliberations. A minimal view of the virtuous agent is of someone who is motivated to perform obligatory acts and avoid performing wrong acts. Insofar, then, as knowledge of what normatively explains why acts are right or wrong can be a means to knowledge of which acts are right or wrong, normative explanations can play an instrumental rôle in helping virtuous agents to achieve their goals of virtuous action.

The connection between normative explanation and virtue that I wish to explore goes deeper than those just mentioned. The *locus classicus* of this position is Kant's analysis of the virtuous, or *good-willed* agent. Kant has traditionally been read as claiming that the virtuous person is the one who is motivated to do what is right because it is right, and to avoid doing what is wrong because it is wrong. Whatever the merits of this reading of Kant, a number of authors have recently defended an amended account according to which the good-willed agent is one who is motivated to act by the presence of features that explain — presumably, *normatively explain* — why certain acts are right or wrong. Korsgaard writes:

Kant is analysing the good will, characterised as one that does what is right because it is right, in order to discover the principle of unconditionally good action. The assumption behind such an analysis is that the reason why a good-willed person does an action, and the reason why the action is right, are the same.⁵²

Following Stratton-Lake, I call this connection between normative explanation the motivation of virtuous agents the *Symmetry* thesis. The essence of the Symmetry thesis is that, in a virtuous person, the motivational explanation of why an action is undertaken coincides with the normative explanation of why it ought to be undertaken. This account of the good-willed agent has a number of advantages over the traditional Kantian account, not least of which is that it allows the virtuous agent to be more sensitively engaged with the underlying features of the moral environment, rather than, as some have seen it, fetishizing moral duty.⁵³ However, Korsgaard presents a raw form of the Symmetry thesis that is open to a number of objections. In response to a number of these objections, Stratton-Lake defends a qualified version of the Symmetry thesis:

A morally good individual is disposed to be motivated to do the right thing by the psychological states which *have as their content* the normative reason why this act is right, and vice versa.⁵⁴

In Appendix A, I will defend a version of the Symmetry thesis from some problems posed for it by Jonathan Dancy's recent work on motivation. For now, however, I will overlook the fine details and rest my argument on the

⁵²Korsgaard 1996, p. 60.

⁵³The fetishizing objection to acting from duty is discussed in *inter alia* Baron 1984.

⁵⁴Stratton-Lake 2000, p. 22 (original emphasis).

weaker view that the Symmetry thesis has a clear and intuitive rationale, and that some version of it is surely correct.

1.4 Conclusion

The central question of my thesis asks after the nature of normative explanations of overall normative facts. In this opening chapter, I have tried to set my project on a sound footing by developing in detail a pluralist theory of explanation, and by characterising in detail my central concept of normative explanation. In so doing, I hope to have made a strong case for thinking that normative explanation is a topic worthy of investigation. By its centrality within ethics, its connections with both first and higher order moral theory, and with virtuous agency, normative explanation provides fertile ground for philosophical investigation.

In Chapter 2, I consider the methodology appropriate to my investigation into the nature of normative explanations, and I clarify some of the subsidiary questions that will be addressed in later chapters. So Chapter 2 sets the detailed agenda for the remainder of this thesis. Now that I have shown in this chapter how the technical aspects of my theory of explanation are to be understood, I will permit myself in later chapters to speak about explanation in terms closer to natural language.

Two

Normative Explanation

2.1 Introduction

The central task of this thesis is to investigate the nature of *normative explanation*, the distinctively normative mode of explanation that I characterized in Chapter 1. In this chapter, I explore a number of issues that are preparatory to my investigation. I begin by considering which methodologies are appropriate to my investigation. I argue that we should begin by investigating the formal properties of normative explanation, and this I do in § 2.2. Next, I offer a minimal defence of Dancy's theory of *enablers* and *disablers*. I then proceed to discuss a number of important but very general issues concerning the nature of normative explanation, some of which are raised by a famous sceptical passage from Mackie. I separate several distinct issues in Mackie's thought and clarify the force of his scepticism. In the final section of this chapter, I explore the possibilities for a reductive account of normative explanation, and I propose a schema for them. I conclude with a brief outline of the arguments to be explored in later

chapters.

2.1.1 Methodology

My central task in this thesis is to investigate the nature of normative explanation, the distinctively normative mode of explanation for facts of overall normative or deontic status. One preparatory task is to consider the methodology appropriate to this investigation. One standard method in philosophical investigation would be to seek a non-trivial account of the truth conditions of normative explanations. This strategy has many virtues, and in Chapters 3–5 I consider a number of applications of this strategy. However, this strategy does have a potential flaw, in that it assumes that normative explanation can be accounted for in other terms. In short, it presupposes the possibility of a *reductive* account of normative explanation. It is highly plausible, however, that not all philosophical phenomena can be reduced to other terms. Some phenomena may be metaphysical primitives, and it cannot be ruled out in advance that normative explanation is among them. But, if normative explanation is a metaphysical primitive, what methodology is appropriate to its study?

Recent work in epistemology provides a model for the study of metaphysical primitives. In *Knowledge and its limits*¹, Williamson argues in detail that knowledge is a metaphysical primitive, and not to be analysed in the traditional terms of justification, truth and belief. Williamson's investigation into the nature of knowledge cannot, therefore, consist in providing a reductive account of knowledge. Instead, Williamson's strategy is to char-

¹Williamson 2000.

acterize knowledge (1) through its formal properties, and (2) through its connections with less primitive phenomena. Let us consider these methods in turn.

Characterising a phenomenon through its formal properties has several virtues. First, studying an entity's formal properties can help to deliver a precise characterisation of it. Second, it can show that our concept of the phenomenon in question is not confused or ill-conceived, and though this may not show that the concept is actually instantiated, it may remove certain objections to the possibility of instantiation. Third, while the search for a reductive analysis presupposes that there is a reductive analysis to be found, one can study an entity's formal properties whether or not it has a reductive analysis. Finally, studying an entity's formal properties has the virtue that, if the entity in question *is* susceptible to reductive analysis, a sound knowledge of its formal properties is helpful in finding the correct form of that analysis.

The second method of investigating the nature of a metaphysical primitive is through its connection with non-primitive phenomena. Roughly speaking, this is the method of reductive analysis in reverse. Instead of defending a phenomenon by reducing it to more primitive metaphysical constituents whose legitimacy is not in doubt, one aims to show that one or more complex phenomena, whose legitimacy is not in doubt, can be reduced to *it* or otherwise explained in its terms. I call this the *constructive* method of analysis, and its explanatory power is increased significantly if the more complex phenomena cannot otherwise be accounted for. However, the constructive method is fully satisfying only in cases where we are dealing with a

genuine metaphysical primitive, so, for this reason, I think the constructive method of analysis *does* presuppose metaphysical primitiveness.

The above considerations lead me to begin my investigation into the nature of normative explanation in § 2.2 by investigating its formal properties. I will then use this formal characterisation of normative explanation to guide discussion of various issues, and the assessment of a number of reductive accounts of its nature. The question of whether the constructive method of analysis is appropriate can be deferred until the various reductive accounts of normative explanation have been considered.

2.2 Formal Properties of Normative Explanation

My aim in this section is to investigate the formal properties of normative explanation. My aim, then, is to investigate the formal properties of the propositional function that is expressed by the connective ‘That p normatively explains that q ’. However, I propose to do this in somewhat indirect fashion. Recall that in § 1.1.4 I proposed a schema for representing the structure of an explanatory connective as a relation between sets of sentences. Letting \mathbb{P} be the set of sentences $\{p_1 \dots p_i\}$ and letting \mathbb{Q} be the set of sentences $\{q_1 \dots q_i\}$, my schema read:

Refined Explanation Schema (RES)

That $p_1 \dots p_i$ and p_i explains that $q_1 \dots q_i$ iff $E(\mathbb{P}, \mathbb{Q})$.

In this schema, an explanatory connective is represented through a relation between the sets of propositions expressed by its constituent sentences. In this way, the formal properties of the *connective* are mirrored in the formal properties of the *relation*. It is helpful to represent connectives in this way because the formal properties of relations are in some ways easier to handle than are the formal properties of functions. I therefore propose to study the formal properties of normative explanatory connective through the indirect, but more perspicuous, means of studying the properties of its associated explanation relation. For the sake of even greater perspicuity I will also drop the set notation in RES whenever it is unnecessary. So, I will often simply write the explanation relation as though its relata were sentences, i.e. $E(p, q)$. In such cases, the set notation remains implicit.

In what follows, I begin with what I take to be the less controversial properties of normative explanation and use those to justify some of my potentially more controversial views. Also, some of the properties I defend may appear to be properties of other modes of explanation besides the normative mode. I will occasionally mention such issues where it serves to clarify my discussion, but a proper defence of the properties of these other modes of explanation claims will usually lie outside the scope of this thesis.

2.2.1 Factiveness

I begin with what I take to be the least controversial claim of all, namely, that normative explanation is *factive*. Indeed, normative explanation is doubly factive: if that p normatively explains that q , then *both* p and q

must be true. Factiveness seems indeed to be a quite general property of explanatory connectives, as illustrated in the truth table for explanatory connectives expressed in Table 1.2. The factiveness of normative explanation is expressed formally as follows:

Factiveness $[E(p, q) \rightarrow [p \wedge q]]$

2.2.2 Irreflexivity

I claim that normative explanation is irreflexive. The irreflexivity of normative explanation can be formally expressed as follows:

Irreflexivity $\neg E(p, p)$

Irreflexivity amounts to the claim that a deontic fact never completely explains itself. For example, an act is never wrong simply because it is wrong. The irreflexivity of normative explanation is most often assumed without argument. Indeed, it is often assumed that irreflexivity is a general feature of explanation. However, Robert Nozick has argued² that there can be reflexive explanations in science, so that assumption may be false. What other arguments can be offered? This question becomes pressing because it has been suggested to me that we can account for the *apparently* irreflexive character of the normative mode of explanation (and perhaps others) by permitting reflexive normative explanation as the trivial case.³ As this picture would have it, normative explanation is like deductive entailment. For, just as every proposition entails itself, but we have good reasons to ignore

²Nozick 1981, Ch. 2, esp. 118–119.

³I owe this suggestion to Brian Ball

circular arguments, so, too, one might claim, every deontic fact explains itself, but circular normative explanations should be ignored on account of their triviality. This is a substantive view, and clearly some argument will be needed to motivate its rejection.

I begin by considering two *epistemic* arguments in favour of irreflexivity. The first epistemic argument is that, if normative explanation were not irreflexive, then there could be cases where normative facts explain themselves, so that, for example, a given act is wrong *because* it is wrong. But, so it might be objected, reflexive explanations of the form ‘*p* because *p*’ should be rejected because they are uninformative.⁴ I return to this argument below.

The second epistemic argument for irreflexivity rests on the *verdictive* character of normative explanations. Normative explanations, so the argument goes, are explanations of why an act has some *overall* or *verdictive* deontic status. *Verdictive* facts deserve the name because, like judicial and other verdicts, they have a certain *epistemic posteriority* to the grounds of those verdicts. More precisely, the idea that overall deontic facts are verdictive implies, not just that overall deontic facts are explained by their grounds, but that it is illegitimate for a verdict to appear among its own grounds. Consider a judicial analogy: the grounds on which a jury reaches its verdict must not include that verdict itself — the verdict is for the jury to decide on the basis of the *other* evidence presented to them. In the same way, so it might be claimed, verdictive normative judgements must be explained on grounds other than themselves.

⁴The argument that reflexive explanations are uninformative is discussed in Stratton-Lake 2000, p. 13. Stratton-Lake attributes the argument to McDowell 1978.

Each of these epistemic arguments has its own faults. In regard to the first argument, one might deny that explanations of the form ‘ p because p ’ are uninformative. Reflexive explanations are not trivial logical truths, for example, and one might claim that they really do provide useful information. They tell us, for example, that the correct explanation of p is p itself, rather than say q or r . As for the second argument, this relies on an analogy with judicial verdicts that can easily be called into question. These arguments also share a common fault, in that each of them seeks to derive metaphysical conclusions from considerations of the nature of normative judgement, and this seems to me to give the wrong direction of fit. I claim that normative judgement should aim to fit the underlying metaphysical facts. So while it may be true that reflexive explanations are uninformative or incompatible with the verdictive nature of moral judgements, whatever constraints — or at least, whatever *non-logical* constraints — there are on normative judgement must ultimately derive from the nature of the underlying phenomena. For this reason, I reject the above epistemic arguments.

For my part, I support the claim that normative explanation is irreflexive with one positive argument and one reply to an objection. My positive argument is that normative explanation is irreflexive because it involves irreflexive normative (or *moral*) laws. I will not defend the claim that normative explanation involves reflexive normative laws here, but, in Chapter 5, I develop a positive account of normative explanation in which irreflexive normative laws play a central rôle.

Consider now an objection to Irreflexivity. Irreflexivity is the claim that deontic facts never wholly explain themselves. Some authors defend the

closely related, but nevertheless distinct, claim that the fact that an act ought overall to be done never provides a *normative reason* for doing it. In recent unpublished work, Parfit objects to this claim on the grounds that if it were true, then, under certain circumstances, an agent could rationally ignore facts about what he ought to do.⁵ This argument from Parfit can be amended to provide a similar objection can be raised against the claim that normative explanation is irreflexive. I will argue, however, that this objection fails.

Let us take the view to be criticised that a deontic fact never normatively explains itself. An amended paraphrase of Parfit's argument⁶ would run as follows (throughout this argument, 'ought' is to mean 'overall ought'): Suppose that some reliable adviser truly tells me that I ought to go home, but does not tell me which facts normatively explain why. If normative explanation is irreflexive, then the fact that I ought to go home does not normatively explain why I ought to go home. But, if that were true, I could rationally decide to stay where I am. I could claim that, though I know that I ought to go home, I am not aware of any fact that normatively explains why I ought to go home. But that would be false (so the Parfitian argument goes). I am aware of the fact that I ought to go home, and this fact clearly explains why I ought to go home. Rather than denying this, our claim should be instead that the fact that I ought to go home only *derivatively* normatively explains why I ought to go home, since its status as a normative explanation derives entirely from the other facts, unknown to me, that *non-*

⁵Parfit 2007, p. 24.

⁶Here I stick as closely as possible to the wording of Parfit's own argument.

derivatively normatively explain why I ought to go home. In summary, the fact that I ought to go home *derivatively* normatively explains why I ought to go home, but its explanatory weight depends entirely on there being some other facts which non-derivatively explain why I ought to go home.

Some might claim that the Parfitian argument poses no real threat to the irreflexivity of normative explanation. All it shows, if successful, is that deontic facts can *derivatively* normatively explain themselves, and it is consistent with this claim that normative explanation is irreflexive at the fundamental level. Congenial though this response may appear, it brings with it a requirement to make sense of the distinction between *fundamental* and *derivative* normative explanation, and this I find difficult. From my point of view, then, it would be preferable if the Parfitian argument could be rejected.

I claim that the Parfitian argument relies on an implicit and false premiss, which we can bring to light in a more formal presentation of the argument:

1. (Premiss) I ought to go home.
2. (Premiss) I know that I ought to go home.
3. (Premiss) I know no fact that fundamentally explains why I ought to go home.
4. (Premiss) It is a rational requirement that (if I know that I ought to go home, then I intend to go home).⁷

⁷Here I assume a *wide-scope* rational requirement, that is, a rational requirement that governs a conditional.

5. (Premiss) For all agents s and for all acts ϕ , if it is a rational requirement that (if s knows that he ought to ϕ , then s intends to ϕ), then *either* s knows some fact which fundamentally explains why he ought to ϕ *or* s knows some fact which derivatively explains why s ought to ϕ .
6. (From 4,5) If it is a rational requirement that (if I know that I ought to go home, then I intend to go home), then *either* I know some fact which fundamentally explains why I ought to go home *or* I know some fact which derivatively explains why I ought to go home.
7. (From 4,6) *Either* I know some fact which fundamentally explains why I ought to go home *or* I know some fact which derivatively explains why I ought to go home.
8. (From 3,7) I know some fact which derivatively explains why I ought to go home.
9. (Premiss) If there is some fact which I know and which derivatively explains why I ought to go home, the fact that I ought to go home is that fact.

Therefore:

10. The fact that I ought to go home derivatively explains why I ought to go home.

The implicit but crucial premiss of Parfit's argument is premiss (5), and I reject it. The Parfitian argument supports premiss (5) with a challenge:

If (5) were false, then I could truly say that I know no fact that explains, whether fundamentally or derivatively, why I ought to go home. But how is this challenge supposed to support premiss (5)? The Parfitian cannot intend that, unless (5) were true, it would not be true that I ought to go home. For it is true, *ex hypothesi*, that I ought to go home (see Premiss 1). Nor can the Parfitian intend that, unless (5) were true, I would not *know* that I ought to go home. For again, *ex hypothesi*, I know that I ought to go home (see Premiss 2). The Parfitian's key claim then concerns rationality. It is the claim that, unless I know some facts which explain why I ought to go home, I would not be irrational if, knowing that I ought to go home, I did not intend to go home. So, the challenge set by the Parfitian argument is to explain how it can be that an agent would be irrational if, knowing that he ought to go home, he did not intend to go home, even though he knows no facts which explain (either fundamentally or derivatively) why he ought to go home. Or, presented another way, the challenge is to explain why I, knowing that I ought to go home, would be irrational if I did not also intend to go home. But, this challenge is easily met: I would be irrational if I did not intend to go home because if I know that I ought to go home, then I *judge* that I ought to go home, and it is irrational to fail to intend to do what one judges that one ought (all things considered) to do. In short, the Parfitian challenge can be answered by a version of *rational judgement internalism*:

Rational Judgement Internalism Necessarily, if s is rational, then, if s judges that he ought to ϕ , s intends to ϕ .

Rational Judgement Internalism is a plausible and widely defended view.⁸ But, note that Rational Judgement Internalism does not support the claim that normative facts explain themselves. Consider a simple example. Suppose I believe that I ought to eat some nearby walnuts, even though, unbeknownst to me, I have a fatal nut allergy. Given my belief that I ought to eat walnuts, I would be irrational if I failed to intend to eat some. But, this cannot explain why I ought to eat some walnuts, because that claim is false. Given my allergy, I ought *not* to eat walnuts. So, the Parfitian argument from irrationality to normative explanation is unsound. With the threat of the Parfitian argument thus defused, the claim that normative explanation is irreflexive may be allowed to stand.

Irreflexivity is the claim that deontic facts never *wholly* explain themselves. However, each of the arguments that I have considered for Irreflexivity implies, in addition, the stronger claim that a deontic fact never *partially* explains itself. To formalize this stronger claim, I cannot avoid recourse to some set notation. The stronger version of irreflexivity can be formalized:

Strengthened Irreflexivity $[E(\mathbb{P}, \mathbb{Q}) \rightarrow \mathbb{Q} \not\subseteq \mathbb{P}]$

In words, the set of propositions in the explanandum of a normative explanation is not a subset of the propositions that make up its explanantia.

2.2.3 Asymmetry

I claim that normative explanation is asymmetric. Formally:

Asymmetry $[E(p, q) \rightarrow \neg E(q, p)]$

⁸Smith 1994, Ch. 3; Wedgwood 2007, Ch. 1.

One might defend the asymmetry of normative explanation as a general feature of all modes of explanation. However, there are persuasive examples of symmetric scientific explanations. For example, suppose we have a physical system consisting of a room in which an energy source powers a thermostatically controlled heater. Let us suppose that this entire physical system is in equilibrium, and will remain so as long as the energy source continues to power the heater. The physical system consists of two subsystems which appear to be in a relation of mutual causation. The heater heats the air and with it the thermostat, thereby causally regulating their temperature. Meanwhile, the temperature of the thermostat causally regulates the output of the electric heater. So we have two subsystems in a relation of mutual causal regulation, and that being so, it is plausible that these two subsystems causally explain each other's behaviour. Hence causal explanation can be symmetric.⁹

If explanation is not in general asymmetric, then some argument is needed to support my claim that *normative* explanation is asymmetric. We might try to defend asymmetry as follows. A normative explanation always involves a non-normative explanans and a normative explanandum. And while non-normative facts may normatively explain normative facts, normative facts cannot normatively explain non-normative facts. Hence, normative explanation is asymmetric. I think that this argument may be only partially correct. That is to say, I accept that there are no normative explanations of non-normative facts. However, I will cautiously suggest in § 2.2.4 and § 2.4.1 that one normative fact may normatively explain another

⁹I am grateful to Dennis Lehmkuhl for helpful discussion on this point.

normative fact. For example, the fact that one act (a promise-breaking) was wrong may explain why another act (making amends) is obligatory. So, asymmetry cannot be defended in precisely this way.

We can improve on the previous argument as follows. Normative facts are normatively explained either by other normative facts or non-normative facts. If they are explained by non-normative facts, then those explanations are asymmetric, because there are no normative explanations of non-normative facts. This argument takes us, as it were, half way to the asymmetry of normative explanation. What remains to be shown is that in no cases do two normative facts normatively explain each other. While I know of no argument that such cases are impossible, they nevertheless seem inconceivable. For this reason, I will assume that normative explanation is asymmetric.

2.2.4 Transitivity and Intransitivity

I turn now to consider whether normative explanation is transitive. Formally, the claim to consider is:

Transitivity $[[E(p, q) \wedge E(q, r)] \rightarrow E(p, r)]$

One might defend transitivity as follows. Normative explanations necessarily have normative *explananda*. So if $E(p, q)$, then q is a normative fact. However, normative explanations *cannot* have normative *explanantia*. So it cannot also be that $E(q, r)$. Hence the antecedent of Transitivity cannot be true. Hence transitivity is trivially true. Note, however, that this argument would also show that normative explanation is trivially *intran-*

sitive, i.e. $[[E(p, q) \wedge E(q, r)] \rightarrow \neg E(p, r)]$. If there is a problem with this argument, it is that one might deny that normative explanations must have non-normative explanantia. To repeat my earlier example, the fact that one act (a promise-breaking) was wrong may seem normatively to explain why another act (making amends) is obligatory. But, if one normative fact can explain another, then neither transitivity nor intransitivity will be trivially true.

Suppose for the sake of argument that normative facts can explain other normative facts. Might normative explanation be non-trivially transitive? Consider the following case. Tim promises Jane that he will take her to the railway station in his car at 12 o' clock, and thereby acquires an obligation to do so. However, Tim fails for no good reason to keep his appointment with Jane, thereby breaking his promise to her. So, Tim acts wrongly in breaking his promise, and, let us suppose, Tim acquires for this reason a new obligation to make amends to Jane for the inconvenience he has caused her. Meanwhile, Jane pays a taxi to the railway station using her own money, so that Tim's obligation to make amends includes at least the obligation to repay Jane's taxi fare. In this case, we have at least the antecedent of a possible instance of transitivity. First, the fact that Tim promised to drive Jane to the station normatively explains why Tim's failure to drive Jane to the station was wrong. Second, the fact that Tim's failure to drive Jane to the station was wrong normatively explains why Tim ought to repay Jane's taxi money. The question to consider then is this: does the fact that Tim promised to drive Jane to the station normatively explain why he ought to repay Jane's taxi money? I claim that it does not. Tim ought to

repay Jane's money because he wrongly broke his promise to drive her to the station. It is not the case that Tim ought to repay Jane's taxi money because he promised to drive her to the station.

In denying this instance of transitivity, I need not deny that Tim's original promise is wholly irrelevant to the normative explanation of why Tim ought to repay Jane's taxi money. Indeed, there is an obvious way in which Tim's original promise is relevant to the normative explanation of why Tim ought to repay Jane's money: Tim's original promise normatively explains that his failing to drive Jane to the station was wrong, and, in turn, this fact normatively explains why Tim ought to repay Jane's money. Thus, Tim's original promise is explanatorily relevant to his obligation to repay Jane's taxi money because it makes an appearance in a chain of normative explanations that leads from the one to the other. But saying this does not commit one to the claim of transitivity. Instead, we can define a wider notion of 'That p appears in the normative explanatory history of the fact that q ', which is the ancestral of normative explanation. Then, we can say that the fact Tim promised to drive Jane to the station appears in the normative explanatory history of the fact that Tim ought to repay Jane's taxi money, without claiming that it normatively explains it.

Let us continue to assume for the sake of argument that one normative fact can explain another. I have argued that, on this assumption, normative explanation is not transitive. Might it be *intransitive*? Formally, the claim to consider is:

Intransitivity $[[E(p, q) \wedge E(q, r)] \rightarrow \neg E(p, r)]$

I know of no general considerations that would support intransitivity. One might perhaps support it inductively, by considering a large number of cases like the case of Tim and Jane, but I shall not pursue an inductive argument here.

My conclusions regarding transitivity are tentative. *If* normative explanations must have non-normative explanantia, then normative explanation is, trivially, both transitive and intransitive. However, *if* normative explanations can have normative explanantia, then normative explanation would seem to be at least non-transitive, and it may be intransitive.

2.2.5 Non-Monotonicity

In some respects, the normative explanation relation resembles deductive entailment. For example, in both relations, one set of propositions grounds a second set of propositions. I have already noted one dissimilarity, namely that normative explanation is factive whereas entailment is not. A second dissimilarity, I suggest, is that normative explanation is *non-monotonic*. A relation is non-monotonic iff it is not monotonic, so let us discuss monotonicity.

In a philosophical context, monotonicity refers to the property of deductive entailment whereby one cannot render a deductively valid argument invalid by, as it were, adding further premisses. Thus, if a set of premisses $\{p_1, \dots, p_i\}$ deductively entails the conclusion that q , then for any proposition r , the set $\{p_1, \dots, p_i, r\}$ also entails that q . For entailment this property holds even if r is false, indeed, even if r is necessarily false.

The same is not true, however, for normative explanation. Suppose that Jim's promise to return John's book normatively explains why Jim ought to return John's book. Now take any further proposition e.g. that grass is green. Is it necessarily true that if the fact that Jim promised to return John's book normatively explains why Jim ought to return John's book, then the fact that Jim promised to return John's book and grass is green normatively explains why Jim ought to return John's book? I claim that this is not a necessary truth, because not true at all. That grass is green is not even a part of the relevant normative explanation. Hence normative explanation is non-monotonic.

The claim of non-monotonicity can be formally expressed as follows:

Non-Monotonicity $\neg \Box \forall \mathbb{P} \forall \mathbb{Q} \forall r [E(\mathbb{P}, \mathbb{Q}) \rightarrow E(\mathbb{P} \cup \{r\}, \mathbb{Q})]$

It might be thought that the non-monotonicity of normative explanation follows directly from its factiveness. But, that is not so. Factiveness takes us part of the way by ensuring that a connection of normative explanation can be undermined by the addition of a *falsehood* to the set of explanantia. But factiveness does not account for the fact that connections of normative explanation can sometimes be undermined by padding the explanantia with extra *truths*.

2.2.6 Intensionality and Extensionality

Consider next the extent to which normative explanations are intensional or extensional. I will argue that normative explanation is neither purely intensional nor purely extensional. Normative explanation is *extensional* insofar

as it provides a referentially transparent context. But, it is *intensional* insofar as it does not permit substitution of logically equivalent sentences. I argue these claims in turn.

I claim that normative explanation is extensional in providing a referentially transparent context that permits the substitution, *salva veritate*, of co-referring expressions. Let $p(a)$ be a sentence containing one or more occurrences of a name 'a', and let $q(b)$ be a sentence containing one or more occurrences of name 'b', and consider the claim that:

(2.2.1) That $p(a)$ normatively explains that $q(b)$.

Now let $p(c)$ be the result of replacing one or more occurrences of the name 'a' in $p(a)$ with a co-referring name 'c', and let $q(d)$ be the result of replacing one or more occurrences of 'b' in $q(b)$ with the co-referring name 'd'. And consider the claim that:

(2.2.2) That $p(c)$ normatively explains that $q(d)$.

If normative explanation is referentially transparent, then 2.2.1 and 2.2.2 will have equivalent truth values. I claim that this is indeed the case with normative explanation. Let us consider a concrete example.

Suppose that there is a cat in my neighbourhood whom I know as 'Tiddles', but whose official name, unbeknownst to me, is 'Montgomery'. Suppose also that:

(2.2.3) That Tiddles is hungry normatively explains why I ought to feed Tiddles.

I claim that 2.2.3 is true iff it is also true that:

(2.2.4) That Montgomery is hungry normatively explains why I ought to feed Montgomery.

I claim that 2.2.3 is true iff 2.2.4 is true because the only fact that seems relevant to my obligation to feed Tiddles is that he is hungry. That is to say, I ought to feed Tiddles because Tiddles, *however he is referred to*, is hungry. Of course, I may *know* that 2.2.3 expresses a truth without *knowing* that 2.2.4 expresses a truth. But, this is consistent with the claim that 2.2.3 is true iff 2.2.4 is true. Furthermore, the referential transparency of normative explanation seems to be a quite general feature of normative explanations. One has obligations because things are thus and so, independently of how they are represented as being thus and so. So, I claim that normative explanation provides a referentially transparent context.

The claim that normative explanation is referentially transparent is highly significant given my claim, in § 1.1.3, that explanatory connectives are not wholly truth-functional. It is significant because there is an important class of arguments, known as *slingshot* arguments, according to which if a connective is both referentially transparent and permits substitution of logical equivalents, then it must be wholly truth functional.¹⁰ Fortunately, I wish to deny that normative explanation permits substitution of logical equivalents, so the slingshot argument poses no threat to my position.

To show that normative explanation does not permit substitution of

¹⁰See esp. Davidson 1967; Davidson 1969; Taylor 1976; Neale 2001; Rodriguez-Pereyra 2003.

logically equivalent sentences, consider an example. Suppose:

(2.2.5) I ought to give a large amount of money to Oxfam because my giving a large amount of money to Oxfam will maximise happiness.

If 2.2.5 is true, then it is true that my giving a large amount of money to Oxfam will maximise happiness. Furthermore, if normative explanation permits substitution of logically equivalent sentences *salva veritate*, then 2.2.5 is true iff it is true that:

(2.2.6) I ought to give a large amount of money to Oxfam because (my giving a large amount of money to Oxfam will maximise happiness and either grass is green or grass is not green).

The parentheses in 2.2.6 delimit the scope of the explanans. However, while 2.2.5 is plausibly true, 2.2.6 is definitely false. So normative explanation does not permit substitution of logically equivalent sentences.

The argument that I have given above has a noteworthy consequence. I have claimed that explanatory connectives do not permit substitution of logical equivalents. That is to say, I deny that $E(p, q)$ entails, and is entailed by, $E([p \wedge [r \vee \neg r]], q)$. However, I maintain that the normative explanatory connective expresses a function from propositions to propositions. So, my argument implies that the proposition that p is not identical to the proposition that $[p \wedge [r \vee \neg r]]$. However, p and $[p \wedge [r \vee \neg r]]$ are true in all the same possible worlds. So, my argument implies that *propositions cannot be identified with sets of possible worlds*. More precisely, the situation here is that the following propositions form an inconsistent set:

1. Explanatory connectives express propositional functions.
2. Explanatory connectives do not permit free substitution of logical equivalents.
3. Propositions are sets of possible worlds.

I have already defended 1-2, so I take there to be good reason to deny that propositions are sets of possible worlds. Moreover, I suggest that rejecting 3 is the most conservative option, the one that requires the least rearrangement of established common sense. If we retained 3 instead of rejecting it, we would need to reject either 1 or 2. 1 is a foundational assumption of my project, so I do not take its rejection to be a possibility. However, to reject 2 would lead, via the slingshot argument, to the claim that normative explanation is wholly truth-functional. This would imply that either every truths normatively explains every truth, or no truths normatively explain any truths. In other words, normative explanation would be ubiquitous or impossible, a result which I take to be catastrophic. I claim, then, that the most conservative strategy is to deny that propositions are sets of possible worlds.

The hybrid nature of normative explanation as neither wholly extensional nor wholly intensional has interesting implications. One interesting implication concerns the possibility of a connection between explanation and understanding. Many philosophers have claimed that explanation and understanding are intimately connected. For example, Strawson writes:

[...] As a first approximation, one could say that the non-natural fact that the explaining relation holds between the fact

that p and the fact that q expands into the natural fact that coming to know that p will tend, in the light of other knowledge (or of theory) to induce a state which we call ‘understanding why q ’.¹¹

Though I am sympathetic to Strawson’s view, the referential transparency of normative explanation means that the details of any such connection will not be straightforward. For example, the explanation that I ought to buy cat food because Tiddles is hungry may well put me in a position to understand why I ought to buy cat food. But if I do not know that Tiddles is Montgomery, the explanation that I ought to buy cat food because *Montgomery* is hungry may not induce a similar level of understanding. So, a full account of the connection between explanation and understanding will need to overcome the obstacles to understanding that referential transparency occasions.

A second interesting consequence of the hybrid nature of normative explanation concerns the possibility of non-objectivist accounts of normative explanation. I postpone discussion of this issue until § 2.4.2.

This completes my discussion of the formal properties of normative explanation. I have argued that normative explanation is factive, asymmetric, and non-monotonic; that it permits substitution of co-referential expressions but not of logical equivalents. I have also argued that *if* normative explanations must have non-normative explanantia, then normative explanation is trivially both transitive and intransitive. Otherwise, it is plausibly non-transitive and may be intransitive. In discussing these formal properties, I hope to have shown that our *concept* of normative explanation is

¹¹Strawson 1985, p. 117.

well-developed and free of obvious confusions. I will now put some of the formal properties I have defended to work in discussing some a number of further general issues.

2.3 Enablers and Disablers

The first general issue to be discussed concerns *enablers* and *disablers*. In recent years, Jonathan Dancy has argued that there are a different ways in which a fact can be relevant to a normative explanation. It is agreed on all sides of this debate that one way in which a fact can be relevant to a normative explanation is by being the whole of, or a part of, its explanans. Dancy's controversial suggestion is that there are other rôles besides this one which are of relevance to normative explanations. Among the rôles distinguished are those of *enabling and disabling conditions*, and *intensifiers and attenuators*. Here I shall be concerned only with enabling and disabling conditions. Dancy's concern with these alternative explanatory rôles originates in the theory of normative reasons, but he extends their use to normative explanations of overall deontic facts:

Just as the favouring relation can be enabled and disabled, so can the rightmaking or ought-making relation. So there are enablers and disablers for rightmaking as well as for favouring.¹²

Much of what Dancy has written on the existence of enablers and disablers relies on the persuasive power of certain well-chosen examples. Here is one of them. Consider a scenario in which all of the following are true:¹³

¹²Dancy 2004*b*, p. 41.

¹³Cf. Dancy 2004*b*, p. 38.

1. I promised to ϕ .
2. My promise to ϕ was not given under duress.
3. I am able to ϕ .
4. There is no greater reason not to ϕ than to ϕ .
5. I ought to ϕ .

The question is how the fact that I ought to ϕ is to be explained. Assume that (1), the fact that I promised to ϕ , is at least *part* of the normative explanation of why I ought to ϕ . It is also plausible that (2)–(4) are in some way relevant to the normative explanation of why I ought to ϕ . The question, then, is whether (1) is the whole of the normative explanans, or only a part of it alongside one or more of (2)–(4). Dancy argues persuasively that (2)–(4) are not part of the explanans. Instead, their rôle is — or at least, it can be in particular cases — to enable the fact that I promised normatively to explain, all by itself, why I ought to ϕ .

Dancy’s theory of enablers and disablers is highly intuitive, but it leaves a good deal of room for interpretation. In a recent paper, Nick Zangwill draws a similar distinction between facts which normatively explains and those which are otherwise relevant to a normative explanation in terms of *responsibility* and *relevance*, and he asks:

But how are we to distinguish responsible from relevant natural properties? How can we know which are right-makers and which are right-allowers? It is one thing to make a principled distinction and another to have a means of telling when it applies.

Furthermore, if we make some abstract philosophical distinction but could not apply it in practice, that would cast doubt on the abstract distinction.¹⁴

One possible response to Zangwill's question runs as follows. The descriptions 'enabling condition' and 'disabling condition' suggest that (2)–(4) play an active rôle in explaining why some fact does or does not normatively explain another fact. For example, one might say that the fact that my promise to ϕ was not given under duress is an *enabling condition* in the sense that it wholly or partially explains why the fact that I promised to ϕ succeeds in normatively explaining why I ought to ϕ . This conception of enablers and disablers suggests that their role is really to be located at the level of *second order* explanations of *first order* normative explanations, a distinction I discussed previously in § 1.3.2.

My task at present is not to develop a positive theory of enablers and disablers, but rather to defend the view that some distinction of this kind is legitimate, by showing the implausibility of denying it. There are two reasons why one might take (2)–(4) to be part of the normative explanation why I ought to ϕ .

- (i) Each of (2)–(4) is part of the explanation of why I ought to ϕ because each is an *indicative necessary condition* of (1)'s at least partially normatively explaining why I ought to ϕ . That is, if (1) explains, at least partially, why I ought to ϕ , then (2)–(4) are true.
- (ii) Each of (2)–(4) is part of the explanation of why I ought to ϕ because each is a *counterfactual necessary condition* of (1)'s at least partially

¹⁴Zangwill 2008, §4.

explaining why I ought to ϕ . That is, if any of (2)–(4) were not the case, it would not be the case that (1) explains, at least partially, why I ought to ϕ .

I discuss each possibility in turn.

2.3.1 Indicative Necessary Conditions

Assume that (2)–(4) in our example are each indicative necessary conditions of the fact that I promised to ϕ providing at least a partial normative explanation of why I ought to ϕ . Should they, for this reason, be counted as part of what explains why I ought to ϕ ? To claim so is to claim that every indicative necessary condition of a set of facts being a normative explanation of a deontic fact is also part of the total normative explanation of that deontic fact. Formally:

$$(2.3.1) \quad \forall \mathbb{P} \forall q \forall r [[E(\mathbb{P}, q) \rightarrow r] \rightarrow r \in \mathbb{P}]$$

2.3.1 states a closure principle for any set of facts that normatively explains an overall deontic fact, and it is very implausible. To illustrate why, let us make the situation more concrete. Let q be the fact that I ought to ϕ , and let \mathbb{P} contain at least the proposition that I promised to ϕ . Then 2.3.1 entails that every indicative necessary condition of the fact that I promised to ϕ being a partial normative explanation of why I ought to ϕ , is also part of the total normative explanation of why I ought to ϕ . This leads to two serious problems.

First, it follows from the factiveness of normative explanation that $E(\mathbb{P}, q)$ has as one indicative necessary condition the truth of q . So 2.3.1 entails that the fact that I ought to ϕ is part of the normative explanation of why I ought to ϕ . However, this contradicts the claim of Strengthened Irreflexivity defended in § 2.2.2. Second, $E(\mathbb{P}, q)$ is, of course, an indicative necessary condition of itself. So 2.3.1 entails that $E(\mathbb{P}, q)$ is itself part of the explanation of q , but this claim is very implausible. Third, *every truth* is an indicative necessary condition of $E(\mathbb{P}, q)$, because, for every true proposition r , $E(\mathbb{P}, q)$ implies r . So, 2.3.1 entails that the total explanation of why I ought to keep my promise includes such contingent truths as that grass is green, or that a big bang occurred, in addition to all the necessary truths, such as those of logic and mathematics. But this result is absurd.

It is clear from these results that 2.3.1 leads quickly to absurdity, so it should be rejected. Not every indicative necessary condition of some fact's partially explaining an obligation should be included alongside it as part of the entire normative explanation of that obligation.

2.3.2 Counterfactual Necessary Conditions

A second possible view is that conditions (2)–(4) in our example must be part of the normative explanation of why I ought to ϕ because for each of them, it is true that, if it were not the case, it would not be the case that the fact that I promised to ϕ is part of the normative explanation of why I ought to ϕ — perhaps because, if that fact were not the case, it would not be true that I ought to ϕ . Letting ' $p \square \rightarrow q$ ' symbolize 'If p were the case, q

would be the case', this view may be stated more formally:

$$(2.3.2) \quad \forall \mathbb{P} \forall q \forall r [[E(\mathbb{P}, q) \wedge [\neg r \sqsupset \neg E(\mathbb{P}, q)]] \rightarrow r \in \mathbb{P}]$$

Like the indicative necessary condition before it, the counterfactual necessary condition in 2.3.2 states a closure principle for the explanantia of normative explanations that is highly implausible. Let us say that when a proposition r is such that $[\neg r \sqsupset \neg E(\mathbb{P}, q)]$, then r is a counterfactual necessary condition of $E(\mathbb{P}, q)$. It follows from the factiveness of normative explanation that q is a counterfactual necessary condition of $E(\mathbb{P}, q)$. So, 2.3.2 entails that q partially explains itself, again contradicting the claim of Strengthened Irreflexivity defended in 2.2.2. $E(\mathbb{P}, q)$ is also a counterfactual necessary condition of itself, so 2.3.2 entails that $E(\mathbb{P}, q)$ is also part of the normative explanation of q . But, this is implausible. Finally, every necessary truth is vacuously a counterfactual necessary condition of $E(\mathbb{P}, q)$, so 2.3.2 entails that every necessary truth is part of the normative explanation of why I ought to keep my promise. But, this result is absurd. These results show 2.3.2 also leads quickly to absurdity.

I take the above arguments to show that there is at least logical space for a theory of enablers and disablers, because to claim otherwise leads quickly to absurdity. Of course, this does not itself provide a positive theory of enablers and disablers. I have suggested that the theory of enablers and disablers should be located at the level of second order normative explanation, because on a natural view, an enabler is a fact which helps to explain why some other fact normatively explains something. I will not develop this view any further here. For my immediate purposes, it suffices to have

shown that the theory of enablers and disablers is plausible, and that there is logical space for it to occupy.

2.4 Mackie's Question

In this section, I draw on work from John Mackie to develop an important sceptical challenge to the possibility of normative explanation. This challenge will be in the background throughout the rest of my thesis, and the argument of my later chapters is structured largely as an effort to respond to Mackie's scepticism. Given its important bearing on my thesis, I will examine Mackie's sceptical challenge in some detail.

When John Mackie argued that moral properties were *queer* he wrote:

Another way of bringing out this queerness it 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? It cannot be an entailment, a logical or semantic necessity. Yet it is not merely that the two features occur together. The wrongness must be somehow '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'?¹⁵

Let us assume that Mackie's focus lies with normative explanations. I will develop Mackie's ideas along two broad lines. The first set of ideas concerns various ways in which deontic facts are dependent on other facts. Here, I will defend Mackie's position but with some restrictions. The second line of thought consists in a sceptical argument from an alleged queerness in the

¹⁵Mackie 1977, p. 41.

nature of normative explanation. I argue that Mackie's sceptical argument is valid, so that the question of its soundness becomes highly significant.

2.4.1 Moral Dependence

Moral properties and moral facts are often said to depend in some way on other properties or facts. Furthermore, this dependence is often thought to be *a priori*. As Jonathan Dancy writes:

We know *a priori* that if an action has a moral property, it has it in view of some other properties which it has.¹⁶

According to some authors, the *a priori* of this knowledge derives from its status as a conceptual truth. Thus, Nick Zangwill writes that the dependence of moral properties on other properties:

[...] is not a just metaphysical constraint on properties but also a constraint on our judgements (a 'conceptual' feature of them). The slogan might be: not just *bad*, but bad because: we judge not that something is bad period, but that it is bad because of certain natural properties. It is *a priori* that moral properties depend.¹⁷

Here I will not be concerned with the alleged status of moral dependence as an *a priori* or conceptual truth. Instead, I wish to isolate a number of non-equivalent moral dependence claims.

Dependency claims can be expressed as relations between individuals, or as connections between propositions. In some cases, this may be a mere notational difference. But, in other cases, for example in truthmaker theory,

¹⁶Dancy 1981, p. 376.

¹⁷Zangwill 2006, p. 271.

the issue is one of substance.¹⁸ In my exposition, I will prefer to express dependency as a connection between propositions. In all of the dependency claims to be discussed, a deontic fact will be taken to depend, in some way or other, on further facts. We can distinguish many different kinds of dependency. I will discuss two broad types of dependency claim for deontic facts: existential and explanatory.

I begin with a very modest version of *existential dependence*. By a *property*, I will mean what is expressed by a monadic predicate; by a *relation* what is expressed by a dyadic predicate. *Overall deontic property* is abbreviated to *deontic property*.

ExtDep 1 Necessarily, for all acts x and all deontic properties D , if Dx then there exists a true non-deontic proposition S .

ExtDep 1 is very modest on several counts. First, it says only that deontic facts depend on some non-deontic facts, though those non-deontic facts might yet be moral facts of some wider kind, such as, evaluative or areteic facts. However, many will wish to defend a stronger claim, viz.:

ExtDep 2 Necessarily, for all acts x and all deontic properties D , if Dx then there exists a true descriptive proposition S .

ExtDep 2 is stronger than ExtDep 1, but it remains modest in other ways. For example, it makes no claim about the logical form of the descriptive proposition S . However, it is plausible that this proposition must state the instantiation of some descriptive property or relation of the act in question.

¹⁸See e.g. Williamson 1999.

Thus:

ExtDep 3 Necessarily, for all acts x and for all deontic properties D , if Dx then either:

1. there exists a descriptive property F such that Fx ; or
2. there exist some entities y_1, \dots, y_i and there exists a descriptive relation R such that x is identical with one of y_1, \dots, y_i and $R(y_1, \dots, y_i)$.

ExtDep 3 expresses a very plausible version of the claim that deontic properties existentially depend on descriptive properties. I claim that ExtDep 3 is true, but not in any interesting way. Consider, for example, *necessary* descriptive properties such as the property of being self-identical, or the property of being such that $2 + 2 = 4$. ExtDep 3 is less interesting for being consistent with deontic properties necessarily implying only *necessary* descriptive properties. A more interesting claim of existential dependence would be:

ExtDep 4 Necessarily, for all acts x and for all deontic properties D , if Dx , then either:

1. there exists a descriptive property F such that it is contingently true that Fx ; or
2. there exist some entities y_1, \dots, y_i and there exists a descriptive relation R such that x is identical with one of y_1, \dots, y_i and it is contingently true that $R(y_1, \dots, y_i)$.

ExtDep 4 expresses perhaps the strongest of the very plausible versions of existential dependence for deontic facts. I claim, however, that its interest is still very limited. For consider: either every deontic property is token-identical with some descriptive property, or it is not. If the former is true, then ExtDep 4 follows trivially. Suppose, then, that deontic properties are not token-identical with any descriptive property. ExtDep 4 is, then, a substantive claim, for it says that deontic properties cannot be instantiated without the instantiation of *distinct, contingent, descriptive* properties. But, this may not tell us anything important about the nature of deontic properties. For, it may be that every entity must have contingent descriptive properties, whether or not it has any deontic properties. For example, Strawson¹⁹ argues that spatial properties, or some analogue of them, are an essential part of the identity conditions of all non-abstract objects. If Strawson is correct, then every act must have spatial properties, and hence, *a fortiori*, every act with a deontic property must have spatial properties. But, this seems to tell us nothing about the nature of deontic properties. There are further examples of the same problem. It is plausible, for example, that every act must have an agent. So, if an act has an overall deontic property, then it also has the descriptive property of being such that there is at least one agent, or the descriptive relation of being done by its agent. But, these claims still seem to tell us nothing significant about the nature of deontic properties.

Claims of existential dependence, substantive though they may be, say nothing about what *normatively explains* deontic facts. For this reason,

¹⁹Strawson 1990.

they are somewhat peripheral to our investigation. Our interest should lie instead with claims of what I will call *normative explanatory dependence*. This is a second family of dependence claims according to which deontic facts depend on other facts in the sense that necessarily, every deontic fact requires a normative explanation. Since normative explanations are factive, *normative explanatory dependence* claims imply their existential counterparts, but they also capture the additional idea of explanatory connection. However, unlike the versions of existential dependence discussed above, almost every claim of normative explanatory dependence is controversial.

I begin again with a weak version of *normative explanatory dependence*:

ExpDep 1 Necessarily, for all acts x and for all deontic properties D , if Dx then there exists a proposition S such that S normatively explains that Dx .

ExpDep 1 is the one of the weakest possible versions of explanatory dependence, but even it can be questioned. More precisely, ExpDep 1 is very plausible for deontic facts concerning *obligation* and *wrongness*. But, it is questionable whether the permissibility of every *permissible* act requires a normative explanation. One reason for denying that permissible acts require a normative explanation is that permissibility may be the default deontic status for acts. It may be, as it were, that all acts are permissible *unless* something normatively explains why they are obligatory or wrong. Pietroski²⁰ defends this view by developing a theory of *prima facie* obligations by analogy with physical forces. The analogy with physical forces leaves room for a *zero force law*, that is, a law that states what happens in

²⁰Pietroski 1993.

the absence of any forces, for which Pietroski proposes the following:

In the absence of *prima facie* obligations, any action is permissible.²¹

Pietroski is concerned with *prima facie* obligations. An analogue of his ‘zero-force law’ for explanations of overall deontic facts might be:

(2.4.1) Necessarily, for all acts x , x is permissible iff there is no proposition S such that either S explains that x is obligatory or S explains that x is wrong.

2.4.1 gives plausible necessary and sufficient conditions for an act’s being overall permissible. But, do they provide a normative explanation of permissibility? Consider whether the following might be true, for some permissible act x :

(2.4.2) The fact that nothing explains why x is obligatory, and nothing explains why x is wrong, normatively explains that x is permissible.

Opinions may divide over whether 2.4.2 is true, but I suggest that it is false. In saying this, I need not claim that there is no mode of explanation in which an action is permissible *because* nothing explains why it is obligatory or wrong. For example, we might that to be a conceptual or ontological explanation, a claim about *what it is* to be permissible. But, the conjunctive fact that nothing normatively explains why an act is obligatory, and nothing

²¹Pietroski 1993, p. 506.

normatively explains why an act is wrong, seems at best to be a non-normative explanation of why x is permissible.

Permissibility, then, may be a special case of deontic property that does not necessarily exhibit explanatory dependency. For this reason, let us put permissibility aside and restrict our attention to obligation and wrongness. Let us call the properties of obligation and wrongness the *special* deontic properties, and the deontic facts that concern obligation and wrongness the *special* deontic facts. ExtDep 1 is very plausible when restricted to special deontic facts.

Consider next whether special deontic facts must depend on descriptive facts. Consider, that is:

ExpDep 2 Necessarily, for all acts x and for all special deontic properties D , if Dx then there exists a descriptive proposition S such that S normatively explains that Dx .

ExpDep 2 is widely assumed. Broad, for instance, writes that:

Moral characteristics are always dependent upon certain other characteristics which can be described in purely neutral non-moral terms. Let us call those non-moral characteristics whose presence in anything confers rightness or wrongness on it right-making and wrong-making characteristics.²²

Recall also that Mackie introduces his question concerning the meaning of ‘because’ by asking how an act’s objective moral qualities are linked with its *natural* features. Similarly, according to Zangwill it is an *a priori*, conceptual truth that moral properties are explained by natural properties: ‘we judge not that something is bad period, but that it is bad because of

²²Broad 1946, p. 103.

certain natural properties.’²³ The difficulty with these view is that there seem to be genuine normative explanations that appeal to evaluative or areteic fact. For example, one might claim that a certain act is obligatory because it brings about the best outcome, or that a certain act is wrong because is unkind or vicious. In § 2.2.4, I also suggested that one normative fact might explain another. In my example, Tim promises to drive Jane to the railway station in his car, and thereby acquires an obligation to do so. But, Tim misses his appointment for no good reason, and thereby does something wrong. I think it conceivable that Tim now acquires a new obligation to make amends to Jane, and that the normative explanation of why Tim ought to make amends to Jane is, precisely, that his act of not driving Jane to the station was wrong. Here, the fact that one act is wrong seem to explain why a second act is obligatory. For these reasons, I think ExpDep 2 is questionable. Some further argument would be needed to accept it.

Consider next a different strengthening of ExpDep 1, according to which:

ExpDep 3 Necessarily, for all acts x and for all special deontic properties D , if Dx then either:

1. there exists a property F such that Fx normatively explains that Dx ; or
2. there exist some entities $y_1, \dots y_i$ and a relation R such that, x is identical with one of $y_1, \dots y_i$ and $R(y_1, \dots y_i)$ normatively explains that Dx .

²³Zangwill 2006, p. 271.

ExpDep 3 claims that every special deontic property that an act has must be explained some property or relation *of that very act*. Note how weak this claim is: it does not even require that the explaining property or relation must be descriptive. Those who would add that further condition will, by implication, be committed to ExpDep 3.

Let us say that a normative explanation of act's having a deontic property is *narrow* iff that act's having that deontic property is normatively explained by *its* having some other properties or relations. ExpDep 3 then amounts to the claim that whenever an act has a special deontic property, there is a *narrow* normative explanation of why it has that deontic property. That deontic properties depend on narrow normative explanations is extremely widely assumed. Recall, for example, Jonathan Dancy's claim that 'We know *a priori* that if an action has a moral property, it has it in view of some other properties which it has.'²⁴ Similarly, W.D. Ross writes:

Tendency to be one's duty may be called a parti-resultant attribute, i.e. one which belongs to an act in virtue of some one component of its nature. *Being* one's duty is a toti-resultant attribute, one which belongs to an act in virtue of its whole nature and of nothing less than this.²⁵

I claim that normative explanations of deontic facts need not be narrow, and in consequence that ExpDep 3 is false. I will offer two reasons for

²⁴Dancy 1981, p. 376.

²⁵Ross 2002, p. 28. It is generally accepted that Ross does not in fact hold the above view. His point here is rather an epistemic one, that one cannot know whether an act is one's duty unless one knows its whole nature, since 'no element in its nature can be dismissed as indifferent.' (Ross 2002, p. 33, footnote 2) Nevertheless, if Ross has only overstated his view, as he claims, rather than grossly misstated it, he must think that an act's special deontic properties necessarily depend on some properties or relations of that very act. Ross' view remains therefore one of narrow normative explanatory dependence.

rejecting the narrowness of normative explanation.

I suggest that much of the motivation for thinking that normative explanations must be narrow comes from the erroneous belief that normative explanations must have a narrow logical form. Certainly, normative explanations are most often formulated in a way that gives them narrow form. But, this narrow form is not inevitable. To show that a wide logical form is possible, consider two possible partial definitions of a crude version of maximizing consequentialism (or *MC*), narrow and wide:

Narrow MC Necessarily, for all acts x , x is obligatory iff the fact that x 's outcome is uniquely best normatively explains why x is obligatory.

Wide MC Necessarily, for all acts x , x is obligatory iff there exists an o such that, o is the outcome of x and the fact that o is uniquely best normatively explains why x is obligatory.

Narrow and Wide MC are partial definitions of maximising consequentialism expressed using narrow and wide normative explanations, respectively. With Narrow MC, an act's causal connection with the best outcome falls inside the scope of normative explanation, and ensures a narrow form of normative explanation. With Wide MC, the normative explanation of why each obligatory act is obligatory is some relation between the act and the best outcome. However, with Wide MC, while the act and its outcome are still related, their relation falls outside the scope of the explanation. Consequently, with Wide MC, the fact that a certain act ought to be done is explained by the fact that its outcome is best. This shows that wide normative explanations are possible.

Wide normative explanations are possible, but are they plausible? In some cases, a wide normative explanation seems better able to capture our normative intuitions. The most obvious examples of this concern what Aristotle classified as *productive* acts. In the opening lines of *Nicomachean ethics*, Aristotle writes:

[...] it is clear that there is some difference between ends: some ends are activities, while others are products which are additional to the activities. In cases where there are ends additional to the activities, the products are by their nature better than the activities.²⁶

The claim that in some cases, the products of an action are ‘by their nature better’ than the acts of producing those products fits better with a wide understanding of normative explanation. In such cases, the normative explanation of why an act ought to be done, seems more properly to concern the properties or relations of its product, and not any relation between the action and its product. But, only wide normative explanations allow us to capture this intuition. Wide-normative explanations also seem to be able to capture part of the idea of what Parfit has called ‘object-given reasons’. Parfit writes:

Object-given reasons are provided by the facts that make certain outcomes worth producing or worth doing for their own sake.²⁷

The idea that outcomes themselves can provide reasons for us to bring them about seems better captured by a wide normative explanation.

²⁶Aristotle 2000, Book 3, 1094a.

²⁷Parfit 2008, p. 44.

To summarise, my argument for the rejection of ExpDep 3 rests on two thoughts. First, there is no barrier from the logical form of normative explanations to allowing this to happen. The assumption that normative explanations *must* be narrow appears to be no more than an artifact of poor logical form. Second, there are cases in which a wide form of normative explanation better captures our intuitions. Together, these claims form a powerful case against ExpDep3.

2.4.2 Mackie's Argument from Explanatory

Queerness

I discussed above a variety of moral dependency claims that arose in the context of my discussion of the famous passage from Mackie cited in § 2.4. That passage from Mackie strongly suggests a sceptical argument, and in this section, I aim to develop it in its best form. The sceptical argument that I aim to develop differs from what are usually called Mackie's *arguments from queerness*. According to Mackie's more familiar arguments, objective moral properties, if they existed, would have to be intrinsically motivating. But, objects Mackie, such properties would be both metaphysically queer, unlike anything else in the universe, and epistemically queer, knowable only by a special epistemic faculty. The argument that I will develop rests on a different source of queerness. So, I call it *Mackie's argument from explanatory queerness*.

Mackie's argument from explanatory queerness has its starting point in the restricted version of ExpDep 1, according to which whenever an act is

obligatory (or wrong), there is some fact that normatively explains why it is obligatory (or wrong). I take it to be highly plausible that obligation and wrongness are dependent in this way. Moreover, ExpDep 1 should be admitted whatever meta-ethical view one holds about the nature of moral judgements. Everyone, whether realist objectivists, realist subjectivists or anti-realist expressivists, should accept this argument's starting point.

Mackie's crucial insight, or so I want to suggest, is that *objectivists* about deontic facts must be *objectivists* about normative explanations of them. With this insight in place, we can now sketch Mackie's argument. Mackie's first premiss is that the special deontic facts necessarily require normative explanations. Next, objective deontic facts require objective normative explanations. But, Mackie suggests, there is no such thing as objective normative explanation. For the objectivist, there is no answer to the question 'what *in the world* is signified by this "because"?' Consequently, there are no objective normative *explanations* of objective deontic facts. So, there are no objective deontic facts. To repeat: if there are any objective deontic facts, then they require objective normative explanations. But, there is no such thing as objective normative explanation. So, there are no objective deontic facts.

We can present Mackie's argument more formally as follows. Let a *special deontic proposition* be a proposition to the effect that a given act is overall obligatory or wrong. Then the argument runs as follows:

Mackie's Argument from Explanatory Queerness

1. (Premiss) Necessarily, for all objective special deontic propositions q , if q then there is an objective normative explanatory function, *explains*, and a proposition p such that p *explains* q .
2. (Premiss) Necessarily, there is no objective normative explanatory function, *explains*.

Therefore:

3. Necessarily, for all objective special deontic propositions q , it is not the case that q .

Mackie's argument presents a beautifully simple and challenging argument to moral objectivism. For those who would deny its conclusion, there are just two premisses to find fault with. However, to deny premiss (1) is to accept that there are occasions when an act is morally obligatory or morally wrong for which there is no normative explanation. In my view, this amounts to a kind of mysticism with respect to morality, so I do not take rejection of premiss (1) to be a live option. This leaves, of course, but one choice. If one accepts that some acts are objectively obligatory or wrong, one must accept that there is such a thing as objective normative explanation. And this of course, brings with the obligation to justify that claim. This is the problem posed to moral objectivism by Mackie's argument: to find an objective account of normative explanation that would underpin claims of the form ' p normatively explains that A ought to ϕ '. I return to this topic in § 2.5.

It would be easy to read into Mackie's argument a potential victory for moral non-objectivism, whether in the form of moral subjectivism or moral non-cognitivism. Commenting on the difficulties his argument poses for the objectivist, Mackie writes:

How much simpler and more comprehensible the situation would be if we could replace the moral quality with some sort of subjective response which could be causally related to the detection of the natural features on which the supposed quality is said to be consequential.²⁸

However, this response may be premature, because there is a general intuition behind Mackie's argument that has implications whatever one's meta-ethical stance. The general intuition behind Mackie's argument is that, whatever meta-ethical stance one takes toward the nature of moral facts, one must take the same stance toward normative explanations of those facts. Objectivists about moral facts must be objectivists about normative explanation; but so, too, subjectivists about moral facts must be subjectivists about normative explanations, and non-cognitivists about moral judgements must be non-cognitivists about normative explanations. So far, I have highlighted the problems that this argument poses for the objectivist. Might it also cause problems for non-objectivists?

One might think that Mackie's argument poses a problem only for objectivism because subjectivism and non-cognitivism have obvious resources with which to run their account of normative explanatory connection. But, consider this issue more closely. Consider a normative explanation of the form '*p* normatively explains *q*'. I assume that both subjectivists and non-

²⁸Mackie 1977, p. 41.

cognitivists will need to give their account of the connection between p and q in this explanation either in terms of a connection between mental states, or in terms of a connection between a mental state and its content. However, in § 2.2.6 I argued that normative explanation provides a referentially transparent context, and this causes a problem because attributions of mental states are paradigm cases of referentially *opaque* contexts. The question therefore arises whether non-objectivists can provide an account of a referentially transparent phenomenon using only referentially opaque resources.

Let me make this problem more concrete with an example. Suppose there is a cat in John's neighbourhood whom he calls 'Tiddles', but whose official name, unbeknownst to John, is 'Montgomery'. Suppose also that:

(2.4.3) John ought to feed Tiddles because Tiddles is hungry.

The subjectivist and the non-cognitivist must try make sense of the normative explanatory connection in 2.4.3 using connections either between mental states or between mental states and their content. Suppose we take a subjectivist line and say that the truth conditions of 2.4.3 consist in the fact that John's belief (or knowledge) that Tiddles is hungry causes John to desire to feed Tiddles. However, if normative explanation provides an extensional context, then 2.4.3 is true iff it is also true that:

(2.4.4) John ought to feed Montgomery because Montgomery is hungry.

This raises two related problems for the subjectivist. The first problem

is that John may know the proposition that Tiddles is hungry under the guise of the sentence ‘Tiddles is hungry’, without knowing that proposition under the guise of the sentence ‘Montgomery is hungry’. The second problem is that John may desire that he feeds Tiddles under the guise of the sentence ‘John feeds Tiddles’, without desiring that he feeds Tiddles under the guise of the sentence ‘John feeds Montgomery’. This is simply a consequence of the fact that propositional attitudes involve intensional rather than extensional contexts. Furthermore, even if John *did* know that Tiddles is hungry under the guise of the sentence ‘Montgomery is hungry’, there is no guarantee that this will cause in him the desire to feeds Tiddles, under any sentential guise, because *ex hypothesi*, John does not know that Montgomery is Tiddles. In consequence, I suggest that the subjectivist has difficulty in account for the fact that 2.4.3 is true iff 2.4.4 is true. The same problem would appear also to arise for non-cognitivism. I assume that a non-cognitivist account of 2.4.3 must invoke either a connection between a mental attitude and its content or a connection between mental attitudes. But, mental attitudes provide intensional contexts, so the non-cognitivist will likewise find it difficult to explain why 2.4.3 is true iff 2.4.4 is true.

The fact that Mackie’s scepticism also causes problems for non-objectivists should not come as any surprise. Mackie accepts that our concepts of moral properties and normative explanation incorporate a presupposition of objectivity, so there must be some elements of ordinary moral thought that non-objectivists are unable to accommodate. One of the ways in which the objectivist nature of moral concepts manifests itself, is in the referential transparency of normative explanation, and the argument above shows

how that can cause problems for the non-objectivist. So, while Mackie's takes his argument from explanatory queerness to pose a special problem for moral objectivism, I claim that it poses a problem for all meta-ethical positions. Nevertheless, I will not pursue these problems for subjectivists and non-cognitivists any further here. My own concern is exclusively with the problems Mackie's argument raises for the objectivist.

2.5 Analysing Normative Explanation

The central question of my thesis concerns the nature of normative explanation. I approach this question from a realist perspective according to which both normative facts, and normative explanations of them, are truth-apt, and some of them are objectively true. However, I have argued in this chapter that this picture of normative explanation faces a serious challenge from Mackie's argument from explanatory queerness. According to this argument, objective facts about obligations and wrongness must have objective normative explanations, but (so Mackie challenges) there is no such thing as objective normative explanation, so there can be no objective normative facts. Mackie's argument provides a beautifully simple and powerful challenge to moral realism, so I propose to orientate the remainder of my thesis as a response to it.

One way to defend the claim that there is such a thing as objective normative explanation is by showing that it is reducible to more familiar terms whose legitimacy is not in doubt. In discussing reductive analyses of normative explanation, I propose to continue to make use of the normative

explanatory *relation* introduced in § 1.1.4 and discussed further in § 2.2. That is to say, I will discuss reductive analyses of the *connective* ‘That p normatively explains that q ’ indirectly by discussing reductive analyses of the normative explanation *relation* ‘ $E(p, q)$ ’.

A correct reductive analysis of normative explanation would, at the very least, provide non-trivial necessary and sufficient conditions for normative explanation, where the requirement of non-triviality implies at least that the analysis is expressed in an alternative vocabulary. Consequently, my aim is to investigate the possibility of finding a non-trivial analysing relation, R , which satisfies these conditions:

Necessity $\square \forall q \forall p [E(p, q) \rightarrow R(p, q)]$

Sufficiency $\square \forall q \forall p [R(p, q) \rightarrow E(p, q)]$

Call these claims the *Necessity* and the *Sufficiency* theses. A candidate analysis must satisfy both if it is to be correct.

2.6 Conclusion

In this chapter, I have developed a better characterisation of normative explanation through some of its formal properties. By showing that we have a well-developed and internally consistent concept of normative explanation, I hope to have given some, albeit weak, support to the view that objective normative explanation is a genuine phenomenon. However, I drew on John Mackie’s work to develop a powerful sceptical challenge to moral realism, resting on a powerful sceptical challenge to the idea of objective normative

explanation. The support provided by a characterization of formal properties seems simply too weak to provide an effective response to this sceptical challenge.

If moral scepticism is to be avoided, an objective account of normative explanation is urgently needed. One natural view, which I developed in Chapter 1, is that explanations involve *necessity*. Accordingly, in Chapters 3–4, I consider attempts to reduce normative explanation to some of the more familiar kinds of necessity, namely, logical, metaphysical, and conceptual. Mackie foresaw these reductive possibilities and claimed that normative explanation ‘[...] cannot be an entailment, a logical or semantic necessity’, but he offered no detailed defence of these claims. Ultimately, however, I concur with Mackie by arguing that these reductive analyses are unsuccessful. In Chapter 5, I aim to develop a possibility that Mackie may have overlooked, by developing an account of normative explanation in which normative laws play a central rôle.

Three

Inferential Accounts of Normative Explanation

3.1 Introduction

In this thesis I seek an account of *normative explanation*, by which I mean a distinctive normative mode of explanation of the overall deontic status of token acts. In this chapter, I investigate whether normative explanation can be reduced to logical necessity, a view which I call *Inferentialism*.

The principal motivation for an Inferentialist account of normative explanation derives from the thought, originating with Aristotle¹, that explanations are *arguments*. The claim that all explanations (of a particular mode) are arguments does not imply the reverse, that all arguments are explanations (of that mode). One established method for narrowing the notion of logical necessity to encompass all and only the explanations of a particular mode is to take explanations to be just those arguments whose

¹See Salmon 1984, Ch. 1; Ruben 1990, Ch. 3.

premisses include non-redundant statements of the *laws* of the domain in question, in our case, moral or normative laws. Inferential accounts of explanation are, for this reason, often called *covering law* accounts.²

There are two principal forms of Inferentialism in the literature. The first and most influential form of Inferentialism is the *deductive-nomological* (hereafter *D-N*) model of explanation, its noteworthy sibling the *Inductive-Statistical* (*I-S*) model. The D-N and I-S models were originally proposed as models of *scientific* explanation, and in that field, the consensus is that each of them is vulnerable to fatal objection.³ In the moral sphere, however, the D-N model of explanation continues to be taken seriously. In this chapter, I will argue that Inferentialism is a poor model for normative explanation, and I will present some new objections which I take to undermine its suitability as a model for *any* mode of explanation. I will argue that D-N arguments are neither sufficient nor necessary for normative explanation. Since every valid D-N argument is also a valid I-S argument⁴, my objections to the sufficiency of D-N arguments count equally against the sufficiency of I-S arguments. For this reason, I will discuss only D-N arguments.

Covering law models of moral explanation have met with trenchant crit-

²Despite the popularity of covering law accounts of explanation, laws are not an essential component of Inferentialism. An Inferentialist account might easily replace laws with, say, generalizations from which the status of *law* is withheld.

³The objections are reviewed in, *inter alia*, Van Fraassen 1980; Salmon 1984; Salmon 1990; Ruben 1990

⁴The I-S model of explanation was developed as a response to the idea that laws of nature may be probabilistic i.e. *statistical*, rather than strict, even at the level of fundamental physics. The I-S model of explanation mirrors the D-N model, but with two principal amendments. First, the conception of laws involved is weakened so as to permit them to be probabilistic rather than strict. Second, the I-S model permits the relation between the explanans and explanandum to be one of *inductive* entailment. D-N arguments can be thought of as I-S arguments in which the laws and the entailment are statistically certain.

icism in recent years, most notably from within the school of *moral particularism*. Particularists charge their opponents with commitment to an objectionable ‘extra-moral view about the nature of *explanation*.’⁵ The most natural candidate for this view is Inferentialism. In this chapter, I, too, will argue against Inferentialism, but my arguments not be particularist.

3.2 D-N Explanation

Contemporary discussion of the D-N model of scientific explanation begins with Hempel and Oppenheim’s “Studies in the Logic of Explanation”.⁶ It is proposed that an explanation (in the sense of a *putative* explanation) consists of two constituents, an *explanandum* and an *explanans*. The *explanandum* is a ‘sentence describing the phenomenon to be explained’. The *explanans* is the ‘class of those sentences which are adduced to account for the phenomenon’. The *explanans* further divides into two sub-classes of sentences, those stating ‘antecedent conditions’, or what we might think of as matters of particular fact, and those representing ‘general laws’.

Four conditions are laid down for a putative explanation to be true (or correct):

1. The *explanans* must logically entail the *explanandum*.
2. The *explanans* must contain at least one general law, and every such general law must be essential, in that the *explanans* would not entail the *explanandum* without it.

⁵Lance and Little 2006*a*, pp. 305-6.

⁶Hempel and Oppenheim 1948. Quotations taken from pp. 136-7 unless otherwise stated.

3. The explanans must have ‘empirical content’.
4. The sentences of the explanans must be true.

The requirement that the explanandum must have ‘empirical’ content is, perhaps, most appropriate to the scientific case, so I propose simply to ignore it here.

The central concepts of the D-N model are *syntactic deductive entailment*, and *law*. I will take syntactic deductive entailment to be sufficiently well understood for our purposes. *Laws*, however, are a different matter. The D-N model of explanation includes the following four conditions on laws.⁷ A *law* is a true *law-like* sentence, where a sentence is law-like iff:

5. It has the logical form of a universally quantified conditional.
6. The domain of its quantifiers is unrestricted.
7. It contains no essential occurrences of designations for particular objects.
8. It contains only *purely qualitative* predicates, where a predicate is purely qualitative iff ‘a statement of its meaning does not require reference to any one particular object or spatio-temporal location.’⁸

In addition, I shall assume that logically trivial conditionals, such as $\forall x[Fx \rightarrow Fx]$, are not to qualify as law-like.

A sentence is law-like, then, only if it has the logical form of a universally quantified and non-trivial conditional. A law-like sentence might be of

⁷Hempel and Oppenheim 1948, pp. 152–7.

⁸Hempel and Oppenheim 1948, p. 156.

narrow form, involving quantification over a single variable e.g. $\forall x[Fx \rightarrow Gx]$, or of *wide* form, involving quantification over several variables e.g. $\forall x[Fx \rightarrow \exists y Gy]$. However, logical form alone is not sufficient to distinguish laws from true generalizations that are merely accidental. For that, conditions (6)–(8) are required. Let us call the conception of laws required for the D-N model, the *strict laws* conception.

Part of the rationale behind conditions (6)–(8) is that they offer ways of distinguishing between *accidental* and *non-accidental* generalizations without invoking either any modal concepts or the concept of explanation itself. The proscription on invoking the very concept of explanation to be analysed is clearly legitimate if the analysis is not to be circular. It is less certain, however, that the proscription on modal concepts is legitimate, because one obvious way of distinguishing between accidental and non-accidental generalizations would be in terms of their metaphysical contingency or necessity.

There were several reasons why the proscription on modal concepts was originally imposed. First, modal notions were still poorly understood in 1948 when Hempel and Oppenheim published their original paper. Second, it was thought that explanation and causation were themselves modal notions, so that any modal analysis of them would be circular. Third, it was thought that there were general ‘empiricist’ grounds for avoiding modal notions in context of scientific explanation. This view might be defended, for example, either on the grounds that scientific method is not able to establish modal truths, or alternatively, because one’s best model of science requires that laws of nature can vary across possible worlds. Whatever the bearing of those considerations on the scientific case, it is arguable that

the same considerations do not carry over to the moral case. Modal logic, for example, is now well developed, and it seems perfectly legitimate to define non-accidental generalizations in terms of their metaphysical necessity. Similarly, many moral philosophers would give some credence to the thought that our methods of establishing moral truths, opaque though they are, are sufficient to establish some of them as metaphysically necessary. For our purposes, then, I will allow that moral laws might take the form of metaphysically necessary universal generalizations. The notion of deductive entailment will consequently need to be that of modal logic rather than simple predicate logic. The question of which modal logics are apposite will be addressed in the next section. I proceed now to the question of whether normative explanation can be analysed in terms of D-N arguments.

3.3 An Inferential Analysis of Normative Explanation

The D-N model of explanation was originally intended as a model of *scientific* explanation. But, it can be generalized to all modes of explanation. So understood, the many *prima facie* distinct modes of explanation, such as causal explanation, conceptual explanation, constitutive explanation, moral explanation, etc., may be distinguished according as the laws involved are causal, logical, constitutive, moral, or what have you. Consider, for example, the prototype D-N explanation that Socrates is mortal by inference from:

1. Socrates is a man; and
2. All men are mortal.

The mode of this explanation is determined by the kind of law expressed in (2). If (2) is a causal or biological law, then we have a causal or biological explanation. For our purposes, then, we shall need to consider D-N arguments involving moral, or more precisely, *deontic* laws. I will take it to be at least a *necessary* condition of a moral law that its consequent concerns a property of overall deontic status, such as being overall obligatory. I noted above that strict laws can take *narrow* or *wide* form. For simplicity, I will focus on laws with a narrow form. Thus, letting D be some predicate of overall deontic status, a moral law must take something like the form $\Box[Fx \rightarrow Dx]$.

I propose to focus on a restricted set of D-N arguments, which I take to provide the standard case for moral explanation. Letting D be any property of overall deontic status, and letting F be any property distinct from D , I propose to focus on D-N arguments with the following schematic form:

Simple D-N Argument Schema

$Fa, \Box\forall x[Fx \rightarrow Dx].$ Therefore Dx .

I will call arguments of this form *Simple D-N Arguments*. Since I am allowing the modal necessity operator, \Box , into the statement of law, the relevant notion of entailment will require a version of modal logic. However, the validity of these Simple D-N Arguments requires only one of the very weakest modal logics, system T. While there are disagreements about which system

of modal logic is correct, those disagreements concern the stronger systems, S4 and S5. All parties to this debate accept the validity of system T, so the validity of Simple D-N Arguments should not be controversial.

I have proposed that we restrict our attention to D-N arguments which fit my Simple D-N Argument Schema. With this restriction in place, the D-N analysis of normative explanation may be stated very simply: sound arguments which fit the Simple D-N Argument Schema are both necessary and sufficient to constitute normative explanations. More formally:

$$(3.3.1) \ E(\{Fa, \Box\forall x[Fx \rightarrow Dx]\}, Da) \leftrightarrow$$

1. Fa ; and
2. $\Box\forall x[Fx \rightarrow Dx]$; and
3. $Fa, \Box\forall x[Fx \rightarrow Dx] \vdash Da$.

Let *D-N Sufficiency* be the thesis that the analysis in 3.3.1 is sufficient for normative explanation. Let *D-N Necessity* be the thesis that it is necessary. I will argue that *D-N Sufficiency* and *D-N Necessity* each face fatal objection.

3.4 Against the D-N Sufficiency Thesis

The claim that D-N arguments are sufficient for *scientific* explanation faces two standard objections, the Asymmetry objection and the Irrelevance objection. In the moral case, I argue that it is debatable whether the D-N

model of normative explanation faces any analogues of the Asymmetry objection, but that there are certainly analogues of the Irrelevance objection.

3.4.1 Asymmetry

The claim that D-N arguments are sufficient for *scientific* explanation is vulnerable to counterexamples of asymmetric deduction. For example, the length of shadow cast by a flagpole can be deduced from the height of the flagpole, the elevation of the sun above the horizon and the laws regarding the propagation of light. But, equally, the elevation of the sun can be deduced from those same laws, the height of the flagpole and the length of the shadow cast by the flagpole, and only the first of these D-N arguments seems genuinely explanatory.⁹

If there are to be analogues of this objection in the moral case, there will need to be cases of symmetric moral laws. I take it to be a necessary condition of a moral law that its consequent involves a moral property. Consequently, if there are to be cases of *symmetric* moral laws, there must be moral laws connecting *two* moral properties. Here is a possible case. Consider the relation between the deontic categories of being *permissible* and being *wrong*. According to a standard view, necessarily an act is permissible *iff* it is not wrong. This seems to provide a case of symmetric moral laws. Necessarily, if an act is permissible then it is not wrong, and necessarily, if an act is not wrong, then it is permissible. Letting '*Px*' abbreviate '*x* is permissible' and '*Wx*' abbreviate '*x* is wrong' one might claim that, then, that there can be pairs of symmetric D-N explanations as follows:

⁹Cf. Bromberger 1966.

$$(3.4.1) \quad Pa, \Box\forall x[Px \rightarrow \neg Wx] \vdash \neg Wa$$

$$(3.4.2) \quad \neg Wa, \Box\forall x[\neg Wx \rightarrow Px] \vdash Pa$$

If these arguments are to provide an Asymmetry objection, one of them must be explanatory and the other not. This is certainly a possible position. One might claim, for example, that permissibility is the most basic deontic category, in whose terms the other categories, including wrongness, is to be explained. So, one might accept that inferences from permissibility to not being wrong are explanatory, but deny that is true of inferences from not being wrong to permissibility. The objection would work equally well if wrongness, rather than permissibility, were the basic deontic category. In either case, it could be claimed that one direction of inference was explanatory, but not the other.

The claim that 3.4.1–3.4.2 provide an Asymmetry objection is vulnerable to two replies. First, one might claim that *neither* direction of inference provides the right kind of *normative* explanation. That an act is not wrong is not to be normatively explained by its being permissible. Nor is an act's being permissible to be normatively explained by its being not wrong. However, if this reply is successful, it would, on its own, mean only that 3.4.1–3.4.2 provide a different kind of counterexample to D-N Sufficiency. If *neither* of 3.4.1–3.4.2 provide a normative explanation, then they would both seem to provide a different kind of counterexample to D-N Sufficiency, a counterexample from Irrelevance.

The second reply to my objection goes further, however, in claiming that 3.4.1–3.4.2 are not counterexamples of any kind to D-N Sufficiency,

because *neither* involves a moral law. According to this second reply, a moral law must have a *descriptive* antecedent and a *normative* consequent. To add this further constraint on moral laws may seem a little *ad hoc*, but it is otherwise plausible. But, if moral laws must take this form, then there can be no Asymmetry objection in the moral case, because there can be no symmetric moral laws. If, for example, it is a moral law that acts with descriptive property *F* are obligatory, it could not also be a *moral* law that obligatory acts have property *F*. For this reason, there may be no Asymmetry objection to the sufficiency of D-N arguments in the moral case. I turn now to the Irrelevance objection.

3.4.2 Irrelevance

According to the Irrelevance objection, D-N arguments are insufficient for normative explanation because they permit irrelevant factors to enter the explanans. I will present two examples.

Consider a moral theory such as Scanlon's contractualism, according to which, necessarily, an act *x* is wrong iff there is no principle permitting *x* that could not be reasonably rejected, by people who were moved to find principles for the general regulation of behaviour that others, similarly motivated, could not reasonably reject.¹⁰ Let '*Cx*' symbolise the property that is necessarily co-extensive with wrongness, and let '*Wx*' symbolise that *x* is wrong. Property *C* meets all the requirements to appear in law-like generalizations, so we have the basis for a first D-N argument:

¹⁰Scanlon 1998, p. 4.

$$(3.4.3) \quad Ca, \Box[Cx \rightarrow Wx] \vdash Wa$$

Now take any property that is necessarily co-extensive with property C , for example, the conjunctive property of being an x such that there are no principles permitting x that could not be reasonably rejected *and* being an x such that there are fewer than seventeen principles permitting x that could not be reasonably rejected. Call this conjunctive property C' . Given that C' is necessarily co-extensive with C , then an act has C iff it has C' . Furthermore, if C necessarily implies wrongness, so, too, does C' . Consequently, given an explanation of the kind in 3.4.3, the D-N analysis entails that there is a second explanation:

$$(3.4.4) \quad C'a, \Box[C'x \rightarrow Wx] \vdash Wa$$

Intuitively, however, contractualists think that there is just one thing that makes acts wrong, namely that there is no principle reasonably permitting them. It is no part of contractualism that wrong acts are wrong partly because there are fewer than seventeen principles reasonably permitting them. The contractualist, therefore, wants to accept 3.4.3 but deny 3.4.4. However, on the D-N analysis of explanation, these two explanations cannot be separated. According to the D-N analysis, these explanations stand or fall together.

One possible reply to my first counterexample is that 3.4.4 represents a harmless form of over-determination. However, I deny this. In § 1.3.1, I argued that normative theories are individuated by the normative explanations they encompass. Consequently, over-determination is not a harmless

phenomenon, because it changes the identity of the theory under consideration. I take it that contractualists would support the idea that 3.4.4 is not a part of their own moral theory, because contractualists, I suggest, will want to *deny* that 3.4.4 provides a genuine explanation of why an act is wrong.

A second reply to this counterexample is that it relies too intimately on the theory of contractualism. However, other counterexamples of this kind can be given using other moral theories. For example, according to one form of consequentialism, necessarily, an act is obligatory iff it is uniquely optimific. But, if actions are necessarily non-abstract objects, then it follows that necessarily, an act is obligatory iff it is uniquely optimific and a non-abstract object. Replacing *C* and *C'* in 3.4.3–3.4.4 with ‘is uniquely optimific’ and ‘is uniquely optimific and a non-abstract object’ would then provide a second counterexample. Again, consequentialists who accept that being uniquely optimific explains why an act is obligatory can reasonably deny that actions are obligatory, in part, because they are non-abstract objects.

I proceed now to a different kind of counterexample. Suppose it to be a necessary truth that God commands an act iff it is morally obligatory. One might object to this as a statement of law since it involves reference to a particular, namely God. For this reason, I stipulate that ‘God commands *x*’ is really to abbreviate ‘It is in the nature of an omniscient, benevolent and omnipotent being to command *x*’, to be symbolized as ‘*Ax*’. Understood in this way, divine command does indeed have a law-like connection with obligation. This makes possible the following D-N argument:

$$(3.4.5) \quad Aa, \Box\forall x[Ax \rightarrow Ox] \vdash Oa$$

If D-N arguments are sufficient for normative explanation, then obligatory acts are obligatory because God commands them, and God's command has a law-like connection to obligation. Many philosophers are persuaded, however, on the grounds of Plato's argument in the *Euthyphro*, that divine command can be a necessary and sufficient condition for obligation without explaining it. If so, then we have a case in which a D-N argument is insufficient for normative explanation.

Some object to the method of testing theories of explanation against our intuitions, and these objections are perfectly reasonable unless our intuitions can be properly developed. In the present case, there are at least two explanations of how divine command might be both necessary and sufficient for obligation without providing a normative explanation of it. The first possible explanation is that there is a moral standard quite independent of divine command, but that God, by his nature, is a perfect tracker of the independently determined truths of moral matters. More precisely, God's perfect benevolence combines with his perfect knowledge to ensure that he commands all and only obligatory acts. The second possible explanation differs from the first in claiming that God *is* relevant to the explanation of moral truths, but only at the level of second order normative explanation. On the second order divine command theory that I introduced in § 1.3.2, God's commanding, on some basis F , an action x , provides a second order, non-normative explanation of why x 's being F normatively explains why x is obligatory. Put another way, divine command is constitutive of first order

normative explanation. Note that *second order* divine command theory is perfectly consistent with the denial of *first order* divine command theory, so long as God commands that we perform acts on some basis other than the fact he commands them. This provides us with a second good account of how divine command can be both necessary and sufficient for obligation, without providing a normative explanation of it.

I have now argued that D-N Sufficiency is vulnerable to the Irrelevance objection, but that it may not be vulnerable to the Asymmetry objection. The Irrelevance objections that I have presented carry straightforwardly over the I-S model of explanation, because every valid D-N argument is also a valid I-S argument. So, every case of a valid D-N argument that is not sufficient for normative explanation, is also a case of a valid I-S argument that is not sufficient for it. Given that the D-N and I-S models are the major models of Inferential explanation, my objections strongly suggest that no Inferential analysis will be sufficient for normative explanation. I turn now to the D-N Necessity thesis.

3.5 Against the D-N Necessity Thesis

According to the D-N model of explanation, a complete explanans comprises a number of particular facts *plus* one or more relevant laws. It bears asking, however, what grounds there are for thinking that the *laws themselves* should appear in the explanans. In this section I will raise four objections to this claim. To be clear, my objection will not be the particularist one that there are no (or at least, no *strict*) moral laws. The target of my objection

is rather the view that the rôle of moral laws in explanation is to appear in the explanans, and thereby make possible a logical inference from explanans to explanandum. My first objection to this claim is that requiring laws to be included the explanans distorts our pre-theoretical intuitions. I accept that this objection only goes so far, but I think it important nonetheless. My second argument is stronger, claiming that the only arguments on offer for the Inferentialist view are either circular or invalid. My third and fourth objections are more ambitious. My third objection aims to show that Inferentialism is internally incoherent. My fourth objection aims to undermine the idea that the D-N is a generally applicable account of explanation. Together, I think these objections form a powerful case against Inferentialism.

3.5.1 The Distortion Objection

According to the D-N model of explanation, a complete explanans necessarily includes at least one relevant law. I object that this requirement distorts our pre-theoretical intuitions.¹¹ Suppose, for example, that I have promised to help John clean out his garage, and thereby acquired an obligation to do so. Pre-theoretically, I want to say that I am obliged to help because I promised to do so. But, if Inferentialism is true, this explanation cannot strictly be accurate, because the explanans contains no relevant law. Or consider the broadly consequentialist view that every obligatory act is obligatory because it maximizes the good. Again, Inferentialism implies that this explanation is never true. This thesis began with the thought that com-

¹¹I owe the seeds of this objection to a discussion with Kit Fine.

nonsense explanations such as ‘I ought to do act *a* because I promised to do so’ or ‘He ought to do act *b* because it will maximize the good’ are in good order, and asking what account could be given of their truth conditions. The first move made by Inferentialism is to deny that these commonsense explanations are in good order, after all. If the Inferential account of explanation is correct, neither the claim that an act is obligatory because it was promised nor the claim that an act is obligatory because it maximizes the good is strictly correct. Each needs supplementing with a relevant law.

There are two sides to the distortion objection. One half of the objection is that including the law in the explanans seems superfluous. Suppose I have promised John to help him clean out his garage, and thereby acquired an obligation to do so. To explain my obligation to help John it seems sufficiently explanatory, all by itself, to mention that I promised to help him. Certainly, there may also be a law to the effect that promises give rise to obligations, but it does not seem *necessary* to mention it in explaining my obligation. The first half of the distortion objection counts against the claim that D-N arguments are *necessary* for explanations.

The second half of the distortion objection is that including the law in the explanans is not merely unnecessary, but falsifying. Consider again my promise to help John clean out his garage. What matters in establishing my obligation, one might insist, is only that I made a promise to John. The fact that there is also a *law* connecting cases of promise-keeping with obligation is, though otherwise an important thing to know, simply not relevant here. This line of objection can be developed further with an

argument I adapt from Alexander Bird.¹² The law that connects promise-keeping with obligations covers all possible cases. We can divide those possible cases into the actual case and the rest. The fact that my promise gives rise to an obligation in the actual case cannot be explained by that part of the law that covers the actual case. For that would be for the actual case to explain itself. But, nor, so it seems, can my obligation in the actual case be explained by the fact that promises imply obligations in *other* possible cases. It is, after all, logically possible that promises should imply obligations in other cases, but not this one. So, the law is in no way explanatory of my present obligation, and to include it in the explanans is to falsify the explanation with irrelevant material.¹³

The distortion objection relies a good deal on appeal to pre-theoretical intuition. This methodology is of course open to two objections, first, that our pre-theoretical intuitions should not be taken to provide the last word, and second, that we, as theoreticians, have no reliable access to our pre-theoretical intuitions. I accept that both of these replies carry some force, so I accept that the distortion objection is not as weighty as it could be. Nevertheless, I think that theoreticians do sometimes have reliable access to pre-theoretical intuitions, both directly, and also indirectly through the moral theories those intuitions have inspired. The moral theory of consequentialism, for example, is often defended as the best way of systematising our pre-theoretical moral intuitions, resulting in the claim that each obligatory act is obligatory because it maximizes the good. The discovery that

¹²Bird 2007, p. 86.

¹³This line of thought also counts against the claim that D-N arguments are *sufficient* for explanation.

Inferentialism is strictly incompatible with a leading form of moral theory should surely give us some pause for thought.

3.5.2 Circularity Objection

My second objection to the view that every explanans must contain a law statement is that the only arguments for it are circular or invalid. I will focus on a recent argument by Roger Crisp, which he gives in the context of a discussion of particularism. Crisp's argument is illustrated with case of causal explanation, but his argument is obviously intended to cover the case of normative explanation. Crisp writes:

As I see it, the generalistic assumption, far from being unmotivated, underlies not only the rationality of ethics, but — on one plausible conception of it — rationality, and its exercise in enquiry.

Consider, again, causal explanation. You may be happy with my explanation of the red billiard ball's moving by reference to its being hit by the white. But now it happens that the white hits the red, and the red fails to move, since it is made of lead. This will lead you to return to the original case: a *fuller explanation* of the move will be that the white hits the red, and that the red is sufficiently light to be moved by the white. [...] It will not do for a rational enquirer to accept that it is just a brute fact that in some collisions, the second ball moves, and in others it does not. To accept such ideas would be deeply irrational [...].¹⁴

I take Crisp to be defending an Inferential account of explanation because of the suggestion that, as one approaches knowledge of a general principle, one

¹⁴34-5 Crisp 2000, (emphasis added).

acquires a ‘fuller explanation’, in the sense of a more complete explanans. Let us consider Crisp’s argument in more detail.

Crisp’s argument is illustrated with a case of causal explanation. A white ball travels across a table, strikes a red ball and causes it to move. So, why did the red ball move? Pre-theoretically, I am inclined to answer that the red ball moved because the white ball hit it. Let us call this the *Naive Pre-Theoretical* view. Inferentialism implies that this could never be a complete explanation, because it lacks a relevant law. We can divide this Inferentialist reply into two components. The first component is that the Naive Pre-Theoretical explanation is incomplete. The second component is that what is required to complete it is a law. Note that these claims are logically independent of each other. I suggest that we should accept the first claim but reject the second.

On the Naive Pre-Theoretical view, the explanation of why the red billiard ball moves is that the white ball hit it. It is a good objection to this view that it lacks much relevant information. One way to develop the Naive view is to accept the requirement that more information is needed, but to claim that all that is required to complete the explanation is more matters of particular fact, not general laws. More precisely, the suggestion is that what is required for an explanation is that there should be some relevant selection of particular facts involved in both the explanans and the explanandum. For example: name the white ball ‘*a*’ and the red ball ‘*b*’. *a* and *b* each have a number of properties, concerning their structure, location, energy states, etc.. The *Improved Pre-Theoretical view* is that there is some subset of *a*’s total properties, which we can abbreviate as

property F , and some subset of b 's total properties, which we can abbreviate as property G , such that b is G because a is F . In this proposed explanation, one agglomeration of particular facts concerning the white ball is taken to explain another agglomeration of particular facts concerning the red ball. The red ball behaves thus and so because the white ball behaves thus and so. We might even add that *which* properties of a explain *which* properties of b is determined by which of their properties are subsumed by the antecedent and consequent of a causal law. Nevertheless, on the current picture, the explanans contains only the particular facts about a that are relevant to explaining the particular facts about b . No law features in the explanans of why the red ball moved.

On the Improved Pre-Theoretical view, an explanans comprises all the matters of particular fact that are relevant to explaining the relevant particular facts in the explanandum. To object to the Improved Pre-Theoretical view on the grounds that it still lacks relevant information now amounts to the claim that there must be some general laws within the explanans. The passage that I quoted from Crisp suggests two arguments for this claim. The first possibility is that an explanans must contain a statement of general law because it must be possible to infer (deductively or inductively) the explanandum from the explanans. But, to offer this argument in support of Inferentialism is clearly question-begging. The second argument that Crisp suggests is that the explanans must include a law because every explanans must relate to its explanandum by *necessity*, rather than their connection being merely brute or accidental. I accept, of course, that explanations require necessity. That is part of the account of explanation that I defended

in Chapter 1. But, it does not follow from this that the connection between explanans and explanandum must be one of logical necessity rather than, say, metaphysical necessity, or a necessity of some other kind (epistemic necessity, for example). So, this second argument is invalid.

Once it is recognised that the necessity that underlies an explanation need not be logical necessity, it becomes otiose to require laws to be a component in an explanations. The D-N model required the inclusion of the law into the explanans only in order to secure a connection of logical necessity. But, if there is no need for the necessity in an explanation to be logical necessity, there is no need for laws to be included in the explanans. Their rôle, if they have one, can be elsewhere.

3.5.3 The Internal Incoherence Objection

My third objection to Inferentialism that it is internally incoherent. My argument rests on a platitude about laws. Laws relate to certain states of affairs by *governing* or *subsuming* them. For example, the law that $\forall x[Fx \rightarrow Gx]$ is in some way supposed to govern or subsume cases in which a given entity is G because it is F . Now, at least part of what it is for a law to govern or subsume a particular case of explanation, is for it to identify all and only the types of fact which are relevant in both the explanans and explanandum. Suppose for example that we have a strict law relating F ness and G ness. And let us suppose that we have an entity, a , that is both F and G . Entity a has, of course, all sorts of other properties besides F and G . And there are also all sorts of other facts in a 's state of affairs. Part of what it is for the law connecting F ness and G ness to subsume a 's case

is for that law to identify that, of all the facts in a 's state of affairs, only a 's being F is relevant to the explanation a 's being G . The platitude that I wish to rely on, then, is that the rôle of laws is to isolate *all, but only* those facts which should appear in the explanans and the explanandum. I will argue that Inferentialism is incompatible with this platitude.

Suppose we start with a naive case in which we suppose it to be true that:

$$(3.5.1) \ a \text{ is } G \text{ because } a \text{ is } F.$$

Suppose also that it is a law that $\Box\forall x[Fx \rightarrow Gx]$. Now, according to Inferentialism, 3.5.1 is strictly false. a is G not simply because a is F . The correct explanation is rather that:

$$(3.5.2) \ a \text{ is } G \text{ because } (a \text{ is } F \text{ and } \Box\forall x[Fx \rightarrow Gx])$$

The round parentheses in 3.5.2 exhibit the scope of the 'because', and so delimit the complete explanans. The problem for the D-N account of explanation now, however, is that the law that is contained within the explanans no longer serves to subsume that explanans. The law that $\Box\forall x[Fx \rightarrow Gx]$ tells us that, the only type of fact relevant to explaining why some entity is G is its being F . But, that is not the only type of fact present in the explanans in 3.5.2. There is also the law that $\Box\forall x[Fx \rightarrow Gx]$. Consequently, there is a mismatch between the explanans and the law in 3.5.2.

One might suppose that this problem is relatively simple to fix. One need only amend the explanans so that it includes a more complex law, as follows:

(3.5.3) a is G because (a is F and $\Box\forall x[[Fx \& \Box\forall x'[Fx' \rightarrow Gx']] \rightarrow Gx]$)

But, this move simply results in a new instance of the same problem. Once again, we are faced with a case in which there are more fact-types in the explanans than its covering law admits as relevant. This problem is bound to recur so long as we try to solve it by building a more complex law into the explanans. I call this the *Subsumption-Relevance* problem.

The only possible solution to this problem is to allow laws to be recursively specified. In our case, for example, that would mean that a 's being G is to be explained as follows:

(3.5.4) a is G because (a is F and L)

where L is recursively defined as the law that: $\Box\forall x[[Fx \& L] \rightarrow Gx]$. The idea that laws *must* be recursively defined poses a deep problem for Inferentialism. For consider. The idea that laws are to be recursively defined is either coherent or incoherent. If it is incoherent, then Inferentialism simply has no answer to the Subsumption-Relevance problem. It might be argued, for example, that it is not possible for laws to be recursively defined. To my knowledge, no such account of laws has been defended, and we would surely need a very good reason to believe that it is coherent. However, even if the claim that laws must be recursively defined is coherent, it is deeply counterintuitive. The claim that laws must be recursively defined is the philosophical equivalent of the epicycles that were needed in Pre-Copernican models of the solar system. Like epicycles, the requirement

that laws be recursively defined may, after all, be a consistent view. But, its consistency comes at such high cost that any other consistent system that supplies equivalent explanatory power without the need for recursively defined laws will necessarily start at a great advantage in terms of credibility.

3.5.4 The Home Defeat Objection

Much of the appeal of Inferentialism stems from its apparent appeal as a generally applicable account of the nature of explanation. One response to the standard objections to an Inferentialist account of scientific explanation, and to the objections I have raised against an Inferentialist account of normative explanation, would be to insist that *some* Inferentialist account of these modes of explanation *must* be correct, and that counterexamples exist to the current Inferentialist accounts only because we have not yet discovered their best form. My final objection seeks to undermine this lingering motivation.

My strategy is to argue that the Inferentialist account of explanation fails on what should be its home territory: logical explanation. I claim that if Inferentialism is to have any plausibility, it must at least work as an account of logical explanation. But, if Inferentialism fails even in that case, then there is no motivation to think that it will work elsewhere.

In 1.2.2 I argued that there is a logical mode of explanation that is exhibited by sound deductive arguments. It is this mode of explanation that underwrites explanations such as:

(3.5.5) B because A and $[A \rightarrow B]$.

Let us apply the D-N model to this logical mode of explanation. As before, the first implication of the D-N model of explanation is that the explanation expressed in 3.5.5 cannot be strictly correct, because the explanans contains no statement of the relevant law of *modus ponens*. The explanation must instead be that:

(3.5.6) B because

1. A ; and
2. $[A \rightarrow B]$; and
3. $\forall p \forall q [[p \& [p \rightarrow q]] \rightarrow q]$.

However, this implication reveals an inconsistency at the heart of the D-N model. Recall that, according to the D-N model, every explanans must contain a relevant law, and every such law must be non-redundant, in the sense that the explanans must not entail the explanandum without it. The case of logical explanation shows how these requirements are in fact inconsistent. The D-N model implies that 3.5.6 is the correct logical explanation of B . However, it will be noted that the law of *modus ponens* stated in (3) is quite redundant to the explanation, because B can already be validly deduced from A and $[A \rightarrow B]$ alone. So, the D-N model requires both that the law of *modus ponens* must be present in the explanans, and also that it must not be present because it is redundant. The D-N model of explanation thus fails as an account of logical explanation. One might reply that logical

explanation is a special case for which an exception is required. But, if logical explanation is not a normal case for the *deductive-nomological* model of explanation, then nothing is. The D-N model of explanation fails precisely where one might have expected it to perform perfectly. And if it fails here, there is no reason to suppose that it will work elsewhere.

3.6 Conclusion

In this chapter I have assessed *Inferentialism*, the claim that normative explanation can be reduced to logical necessity. Inferentialism, in the form of the D-N and I-S models of explanation, is already widely rejected as a model of scientific explanation. In this chapter. I have argued that they fare no better in the moral case, and that such accounts are neither sufficient nor necessary for normative explanation. My most ambitious arguments were designed to show that Inferentialism is in fact not just false, but internally incoherent. I take this to be a new argument in the debate regarding Inferential models of explanation.

I noted in my introduction that covering law models of explanation have met with trenchant criticism from within the school of *moral particularism*. In this chapter, I, too, have argued against Inferentialism, but my arguments have not been particularist. Indeed, I would suggest that, by concentrating their objections on the possibility of moral *laws*, particularists have largely overlooked the specifically *inferential* character of Inferentialism that is one major source of its problems.

Four

Metaphysical and Conceptual Accounts of Normative Explanation

4.1 Introduction

In Chapter 3, I considered and rejected *Inferentialism*, the thesis that normative explanation can be reduced to logical necessity. One deep problem with Inferentialism was the rôle it assigned to moral laws. On the Inferentialist view, moral laws must be components in the explanans of normative explanations, because only then will the complete explanans logically entail the explanandum. The failure of the Inferentialist view frees us to search for an alternative account that avoids this mistake. A viable alternative to Inferentialism should not only analyse normative explanation in terms of a non-logical form of necessity. It should also deny that moral laws are components in the explanantia of normative explanations. If moral laws have a rôle to play in normative explanation, it is elsewhere.

In this chapter, I assess two accounts of normative explanation in which

moral laws can play a different rôle. First, I assess *Metaphysicalism*, the thesis that normative explanation can be analysed in terms of metaphysical necessity. I consider two Metaphysicalist accounts, one in terms of metaphysically necessary implication, and one in terms of supervenience. I argue that both versions fail for similar reasons. Then I discuss *Conceptualism*, the thesis that normative explanation can be analysed in terms of conceptual necessity. I assess two Conceptualist accounts, and I argue that both versions fail, though for different reasons.

I discuss Metaphysicalism and Conceptualism together because some versions of Conceptualism are parasitic on one or other version of Metaphysicalism. Once the relevant Metaphysicalist account is shown to fail, it is a small step to show the failure of any version of Conceptualism that relies on it.

4.2 Metaphysicalism

Metaphysicalism is the thesis that normative explanation can be analysed in terms of metaphysical necessity. I assess two forms of Metaphysicalism in this chapter. The first form aims to reduce normative explanation to metaphysically necessary implication. Call this the *Metaphysical Necessitation* analysis. The second form aims at a more sophisticated reduction of normative explanation in terms of *supervenience*. Call this the *Supervenience Analysis*. One motivation to assess these claims is their popularity. I give examples below of explicit defenders of these views. But, I suggest that many others are implicitly committed to some version of them. The

evidence for this implicit belief lies in the established method of defining first order moral theories. First order moral theories are most often defined through propositions of the following schematic form, where D is some property of overall deontic status:

Biconditional Necessarily, for all acts a : act a is D if and only if . . .

Instances of this schema yield metaphysically necessary biconditionals i.e. propositions which state necessary and sufficient conditions for an act to be e.g. obligatory, that hold in all metaphysically possible worlds. To give just one example, Brad Hooker defines his favoured version of rule consequentialism as follows (the claim of metaphysical necessity is, of course, implicit):

RULE-CONSEQUENTIALISM. An act is wrong if and only if it is forbidden by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being (with some priority for the worst off).¹

Recall, however, that I argued in 1.3.1 that a definition of a first order moral theory is a set of propositions about what explains why acts are right and wrong. Consequently, when instances of the Biconditional schema are used to define a normative theory, it is reasonable to impute to their proponents the belief that they are both necessary and sufficient for normative explanation.

I proceed now to discuss my first version of Metaphysicalism, the Metaphysical Necessitation Analysis.

¹Hooker 2000, p. 32.

4.3 The Metaphysical Necessitation

Analysis

The first version of Metaphysicalism to consider is the *Metaphysical Necessitation* analysis. Let a *strict conditional* be a proposition of the form $\Box\forall x[Fx \rightarrow Gx]$, where the ‘ \Box ’ indicates metaphysical necessity. The simplest version of the Metaphysical Necessitation analysis is:

(4.3.1) Strict Conditional Analysis

$$E(Fx, Dx) \text{ iff } Fx \text{ and } Dx \text{ and } \Box\forall z[Fz \rightarrow Dz].$$

Including Fx and Dx into the analysis ensures that it is factive, while the claim that $\Box\forall z[Fz \rightarrow Dz]$ supplies the component of metaphysical necessitation. Strictly speaking, it is redundant to include Dx alongside Fx and $\Box\forall z[Fz \rightarrow Dz]$, but its inclusion adds clarity at no cost.

4.3.1 represents a crude version of Metaphysicalism that few would defend. For example, it is trivially true that $\Box\forall z[Dz \rightarrow Dz]$, so 4.3.1 is compatible with reflexive normative explanations of the form ‘ x is obligatory because x is obligatory’. I have argued, however, that reflexive normative explanations are illegitimate. Consider, then, two strengthening amendments to 4.3.1. The first amendment involves replacing the strict one-way conditional with a strict *biconditional* i.e. a proposition of the form $\Box\forall z[Fz \leftrightarrow Dz]$. The second amendment involves taking this strict biconditional, or rather, its left-to-right component, to meet the conditions of law-likeness discussed in § 3.2. To recall: our current conception of moral laws is as follows:

(4.3.2) A sentence states a law iff it is metaphysically necessary and it meets the following conditions of law-likeness:

1. It has the form of a universally quantified conditional, whose consequent involves a deontic property.
2. The domain of its quantifiers is unrestricted.
3. It contains no essential occurrences of designations for particular objects.
4. It contains only purely qualitative predicates, where a predicate is purely qualitative iff ‘a statement of its meaning does not require reference to any one particular object or spatio-temporal location.’²
5. It is not a trivial logical truth

On this conception, laws always take the form of a strict one-way conditional i.e. $\Box\forall x[Fx \rightarrow Dx]$, but this account can be extended to cover biconditionals. I will say that a biconditional of the form $\Box\forall x[Fx \leftrightarrow Dx]$ is a *biconditional law* iff it is true and it is a law that $\Box\forall x[F \rightarrow Dx]$. Simply put, a biconditional law is the conjunction of a law of the form $\Box\forall x[Fx \rightarrow Dx]$ with the reverse strict conditional $\Box\forall x[Dx \rightarrow Fx]$.

Strengthening the Strict Conditional Analysis in this way yields:

²Hempel and Oppenheim 1948, p. 156.

(4.3.3) Biconditional Law Analysis

$E(Fx, Dx)$ iff

1. Fx and Dx ; and
2. $\Box\forall z[Fz \leftrightarrow Dz]$; and
3. it is a law that $\Box\forall z[Fz \rightarrow Dz]$.

4.3.3 has one clear virtue over Inferentialism, namely, that it does not take moral laws to be components in the explanans of a normative explanation. Rather, moral laws are themselves taken to underwrite the necessary connection between explanans and explanandum. This point recommends Metaphysicalism over Inferentialism.

I have now presented two forms of the Metaphysical Necessitation analysis of normative explanation, the Strict Conditional analysis in 4.3.1, and Biconditional Law analysis in 4.3.3. For each analysis, we can ask whether it is sufficient or necessary for normative explanation. This gives me four theses to consider, but I need not consider them individually. Note that the Biconditional Law analysis is logically stronger than the Strict Conditional analysis. I will argue that the Biconditional Law analysis is *insufficient* for normative explanation. If that is so, the logically weaker Strict Conditional analysis is also insufficient.

My main task in this section is to argue that no version of the Metaphysical Necessitation analysis, neither the Biconditional Law nor the Strict Conditional analysis, is sufficient for normative explanation. If my arguments are successful, they show that no analysis of normative explanation

in terms of Metaphysical Necessitation is possible. However, there remains the question of whether these analyses are necessary for normative explanation. I will not investigate this question in detail, but I will briefly suggest some reasons for thinking that they are unnecessary.

Before I proceed to criticize the Biconditional Law analysis, let me note some of its explicit defences in the literature. Fred Feldman writes:

A criterion of morality, or fundamental principle of morality, is a statement of necessary and sufficient conditions for obligatoriness, permittedness, and forbiddenness.

[...] if we find a condition, *C*, that is both necessary and sufficient for rightness, then we will have found a *criterion of right action*. We have found a condition, *C*, such that all right acts satisfy *C*, and all acts that satisfy *C* are right. In other words, the class of right acts coincides exactly with the class of acts that satisfy *C*. Then we can say something of this form:

NS : An act is right if and only if it satisfies *C*.

The fundamental principle of morality is some statement of this form.³

I understand Feldman to be claiming that biconditional laws capture claims about what normatively explains why acts are right and wrong. After all, anything as important as the *fundamental principle of morality* should tell us not only *which* acts are right and wrong, but also *why*. So, I take Feldman to be committed to the Biconditional Law analysis of normative explanation.

A second author who defends the Biconditional Law analysis is Mark Timmons. Timmons defines moral principles as:

³Feldman 1978, p. 20.

[...] very general moral statements that purport to set forth conditions under which an action is right or wrong or something is good or bad.⁴

Timmons' moral principles thus take the form of biconditional laws. He continues:

In attempting to satisfy the theoretical aim of explaining what makes an action right or wrong or what makes something good or bad, moral philosophers have typically sought to formulate moral principles that express this information. In fulfilling this theoretical aim, then, a moral principle concerned with right and wrong action can be understood as indicating those most basic features of actions that make them right or wrong.⁵

Here Timmons explicitly claims that moral principles, which have the form of Biconditional Laws, express general information about what normatively explains why acts are right and wrong. The Biconditional Law analysis in 4.3.3 simply applies that general information to token right and wrong acts.

Based on these quotations, I take Feldman and Timmons to be explicitly committed to 4.3.3, or if not exactly that, then something very close to it. At the very least, I take them to be committed to the claim that the Biconditional Law analysis is *sufficient* for normative explanation. If moral principles, which have the form of biconditional laws, were not sufficient for normative explanation, they could not be understood, as Timmons claims they can, as 'indicating those most basic features of actions that make them right or wrong'.

⁴Timmons 2002, p. 5.

⁵Timmons 2002, p. 6.

Consider now G.E. Moore's attempt to define utilitarianism in his *Ethics*⁶. Moore offers a definition of utilitarianism with two official parts.⁷ The first part states that a voluntary act is right iff it produces a maximum of pleasure, and wrong iff it does not. However, Moore goes on to claim that these claims alone would be inadequate for a definition:

Our theory does then, even in its first part, profess to give us an absolutely universal *criterion* of right and wrong; and similarly an absolutely universal *criterion* of what ought or ought not to be done.

But though it does this, there is something else which it does not do. It only asserts, in this first part, that the producing of a maximum of pleasure is a characteristic, which did and will belong, *as a matter of fact*, to all right, voluntary actions (actual or possible), and only to right ones; it does not, in its first part, go on to assert that it is *because* they possess this characteristic that such actions are right. This second assertion is the first which it goes on to make in its second part; and everybody can see, I think, that there is an important difference between the two.⁸

Moore's two-stage definition of utilitarianism has a first part which states that producing a maximum of pleasure is both necessary and sufficient for an act to be right, and failing to do so necessary and sufficient for an act to be wrong. But, it also has a second part which states that it is producing or failing to produce a maximum of pleasure which *explains why* acts are right or wrong. How are the two parts of Moore's definition connected? Unlike Feldman and Timmons, Moore would appear to deny that the biconditional law, connecting obligation with producing a maximum of pleasure,

⁶Moore 1912.

⁷Moore's definition has a third, unofficial part, to be discussed in § 4.4.

⁸Moore 1912, p. 21. Unlike Feldman, Moore uses the phrase, 'criterion of right and wrong' to mean an epistemic criterion. (See Moore 1912, p. 22)

is sufficient to explain why obligatory acts are obligatory. Yet, Moore does not regard the first stage of his definition as dispensable or irrelevant. On the contrary, Moore thinks it indispensable. Consequently, I suggest that Moore takes the Biconditional Law analysis to be necessary for normative explanation.

I proceed now to criticize Metaphysical Necessitation.

4.3.1 Against the Biconditional Law Sufficiency Thesis

Let the *Biconditional Law Sufficiency Thesis* be the claim that:

(4.3.4) $E(Fx, Dx)$ if

1. Fx and Dx ; and
2. $\Box\forall z[Fz \leftrightarrow Dz]$;
3. it is a law that $\Box\forall z[Fz \rightarrow Dz]$.

I will argue that 4.3.4 is false.

Let ' Gx ' symbolise 'God commands x ', and let ' Ox ' symbolise ' x is obligatory'. Consider the claim the God's commanding act x is both necessary and sufficient for x to be obligatory, in symbols:

(4.3.5) $\Box\forall x[Gx \leftrightarrow Ox]$

4.3.5 is equivalent to the following conjunction:

(4.3.6) $[\Box\forall x[Gx \rightarrow Ox] \wedge \Box\forall x[Ox \rightarrow Gx]]$

4.3.5 thus meets all the syntactic conditions to be a biconditional law.⁹ So it would be a moral law, if true. Furthermore, it is widely accepted that the truth of 4.3.5 is at least a conceptual possibility, and for my purposes, I think that is enough.

Assume for the sake of argument that 4.3.5 is true. 4.3.4 and 4.3.5 together imply that divine command is not only necessarily co-extensive with obligation, but that divine command normatively explains why obligatory acts are obligatory. However, as I argued in the last chapter (§ 3.4.2), we can legitimately deny that divine command normatively explains our obligations. For example, it may be that God commands all and only obligatory acts because he is an epistemically and morally ideal observer of the (independent of Him) truths in moral matters. Or, it may be that divine command does have a part to play in explaining why acts are right and wrong, but only at the second order level. God's commanding, on basis F , that act x be done, may constitutively explain why the fact that Fx normatively explains why x is obligatory. Therefore, so long as God commands that we do certain acts because e.g. they help the needy, rather than because of their status as divinely commanded, divine command need have no place within first order normative explanations. This yields a first counterexample to 4.3.4

For a second counterexample, we can take note of some of the logical properties of 4.3.4, and its attendant conception of moral laws. Consider a biconditional law $\Box\forall x[Fx \leftrightarrow Dx]$. If Fx necessarily implies Dx , then any

⁹If reference to God, as a particular, is thought to invalidate the claim of law-likeness, we can replace 'God commands x ' with 'It is in the nature of an omniscient, benevolent being to command x '.

strengthening of property F , such as the conjunctive property of being F -and- H , will also necessarily imply Dx .¹⁰ On its own this does not threaten 4.3.4. But, if F can be strengthened to produce a new property which remains *necessarily co-extensive* with D , this will yield a second biconditional law that is entailed by the first. This result provides a further source of counterexamples to 4.3.4. Let me illustrate.

Take any plausible first order moral theory that is compatible with the Biconditional Law analysis of normative explanation. For the sake of illustration, I will adopt a version of Scanlon's contractualism.¹¹ Let ' Cx ' symbolise 'there is no principle permitting x that could not be reasonably rejected, by people who were moved to find principles for the general regulation of behaviour that others, similarly motivated, could not reasonably reject'. Property C meets all the requirements to appear in law-like generalizations. Let ' Wx ' symbolise that x is wrong. Our contractualism is defined:

$$(4.3.7) \quad \Box \forall x [Wx \leftrightarrow E(Cx, Wx)]$$

Since normative explanation is factive, 4.3.7 entails the biconditional law:

$$(4.3.8) \quad \Box \forall x [Cx \leftrightarrow Wx]$$

4.3.8 combines with the Biconditional Analysis to yield the conclusion that wrong acts are wrong because there is no principle that reasonably permits them. So far, this is consistent with our contractualism, as defined in 4.3.7.

¹⁰Cf. Wedgwood 2007, p. 151.

¹¹Scanlon 1998, p. 4.

Now consider the following conjunctive properties:

1. Mx : Cx and x is such that $2 + 2 = 4$
2. Ix : Cx and x is self-identical.
3. Ax : Cx and x is an act.
4. $\neg Ox$: Cx and x is not obligatory.
5. Sx : Cx and x has spatio-temporal location.
6. $C'x$: there are fewer than seventeen principles permitting x that cannot be reasonably rejected.

These conjunctive properties have some notable features. First, the second property in each conjunction is fitted to appear in law-like generalizations.¹² So, its conjunction with property C is also fitted to appear in a law-like generalization. Second, it is highly plausible that each of these conjunctive properties is necessarily co-extensive with property C alone. For example, an act x has the contractualist property of being such that there is no principle reasonably permitting it iff it has the conjunctive property of being such that there is no principle reasonably permitting it *and* being an act. And if acts necessarily have spatio-temporal location, then an act has the contractualist property iff it has the conjunction of the contractualist property and the property of having spatio-temporal location. Parallel reasoning applies to the properties in 3–6.

¹²One might object that the property of being such that $2 + 2 = 4$ is not fitted to be part of a law since it makes reference to individuals, namely numbers. But, we can remove this difficulty by using instead the property of being such that every even prime is less than every smallest odd prime. This amounts to the property of being such that $2 < 3$, but it is expressed without essential reference to particulars.

I noted above that our conjunctive properties are fitted to appear in laws, and that each of them is necessarily co-extensive with the contractualist property *C*. 4.3.8 thus entails several new biconditional laws, one for each of the conjunctive properties listed above. For example, 4.3.8 entails the biconditional law that necessarily, an act is wrong iff it has the property of being such that there is no principle reasonably permitting it and such that $2 + 2 = 4$. And this, in turn, ensures that the Biconditional Law analysis has false instances. Contractualists want to claim that wrong acts are wrong *just* because there are no principles reasonably permitting them. If the Biconditional Analysis is sufficient for normative explanation, they must also say that wrong acts are wrong because of each of the conjunctive properties listed above. The Biconditional Sufficiency Thesis combines with contractualism to entail that wrong acts are wrong, in part, because they are acts, or because they have spatio-temporal location. But, these claims are implausible. If there is a kind of normative explanation in which these facts play a part, it is a much looser kind than we are looking for.¹³

Consider four objections to my second kind of counterexample. First, one might claim that my first counterexample involves what is only a harmless form of over-determination. However, as I argued in the last chapter (§ 3.4.2), I deny that over-determination is harmless. First order moral

¹³The argument I have just given relies presupposes that propositions are not sets of possible worlds. I claim e.g. that the proposition that a given act is obligatory because it has the contractualist property, and the proposition that it is obligatory because it has the contractualist property and is self-identical, can differ in truth value, and hence are distinct propositions. However, these two propositions are necessarily equivalent, and hence, true in precisely the same possible worlds. My argument thus presupposes that propositions are not sets of possible worlds. My argument for that claim is given in § 2.2.6.

theory are defined by which normative explanations they accept. To accept a second level of normative explanation is to change the theory under consideration.

The second objection one might raise is that my examples illegitimately rely on taking conjunction to be a valid property formation operation. My examples do rely on that, but I can defend my method in two ways. First, conjunction is widely assumed to be a valid method of property formation. Negative properties and disjunctive properties have received serious criticism, but conjunctive properties have been widely, if not universally, accepted.¹⁴ There is, therefore, a serious burden of proof to be met before conjunctive properties should be disallowed. Second, and more importantly, although my counterexamples rely on forming properties through conjunction, they do not *essentially* rely on it. I have presented my counterexamples using conjunctive properties because I am operating under the simplifying assumption that biconditional moral laws take a simple form such as $\Box\forall x[Fx \leftrightarrow Dx]$. This led me to formulate my counterexamples using conjunctive properties, resulting in laws of the form $\Box\forall x[x \text{ is } F\text{-and-}G \leftrightarrow Dx]$. But, this simplifying assumption can be relaxed as needed. It is consistent with the conception of laws expressed in 4.3.2 that they should take a more complex form, such as $\Box\forall x[[Fx \wedge Gx] \leftrightarrow Dx]$. So, instead of objecting that 4.3.3 combines with contractualism to entail that an act x is wrong iff x it has the conjunctive property F -and- G , my objection could be reformulated as the claim that 4.3.4 combines with contractualism to entail

¹⁴Van Cleve (1990) objects to negative properties, while Armstrong (1978) objects to disjunctive properties, but both cautiously accept conjunctive properties. Objections to conjunctive properties are raised by Kraemer 1977; Casullo September 1984.

that an act x is wrong iff x is F and x is G . So, my counterexamples do not essentially rely on taking conjunction to be a valid property formation operation.

Third, one might object to my second kind of counterexample on the grounds that moral laws should not make reference to properties which are necessarily had by every entity, such as the property of being such that $2 + 2 = 4$. Or, if my objection is expressed in terms of conjunctions of facts rather than conjunctive properties, one might object that moral laws should not make reference to necessary truths. I have two replies to this objection, one critical and one concessive. My critical reply is that mathematical truths often appear relevant to normative explanations. For example, suppose one has to share a cake between six people, and one ought to cut the cake at 60 degree angles. It seems to me that the correct explanation of why one ought to cut the cake just so might legitimately make reference to the fact that $360/60 = 6$. So, one might claim that some mathematical truths are relevant in some normative explanations. (My objection to 4.3.3 was not that it made explanation by mathematical truths possible, but that it made it *ubiquitous*.)

But, suppose my critical reply is inadequate, and that one can motivate a general ban on allowing metaphysically necessary truths into moral laws. No matter; the same problem arises with a number of contingent properties. Consider again our representative biconditional moral law $\Box\forall x[Fx \leftrightarrow Dx]$. Let H be the property formed by conjoining F with any other property, I that is necessarily implied by F , but distinct from D . It follows that H is necessarily co-extensive with F , and so, by transitivity of necessary

co-extension, H is also necessarily co-extensive with D . So long, therefore, as I meets the condition of being purely qualitative, the biconditional law that $\Box\forall x[Fx \leftrightarrow Dx]$ entails a second biconditional law, namely $\Box\forall x[H \leftrightarrow Dx]$. Here are two examples of this problem in action. Consider again our contractualist claim that necessarily, an act x is wrong iff there is no principle permitting act x that could not be reasonably rejected. If there is no principle permitting act x that could not be reasonably rejected then there are, *a fortiori*, fewer than seventeen principles permitting x that could not be reasonably rejected. Thus, the biconditional law that $\Box\forall x[\text{there is no principle permitting } x \text{ that could not be reasonably rejected iff } x \text{ is wrong}]$ entails the second biconditional law that $\Box\forall x[x \text{ is such that there is no principle permitting } x \text{ that could not be reasonably rejected and such that there are fewer than seventeen principles permitting } x \text{ that could not be reasonably rejected iff } x \text{ is wrong}]$. This biconditional law relies on no necessary properties, yet it still provides a counterexample to 4.3.4. The fact that there are fewer than seventeen principles permitting an action that could not be reasonably rejected is not part of any explanation of why it is wrong.

Fourth, one might object to my second kind of counterexample on the grounds that it relies on the assumption of a contractualist moral theory. I deny this charge. My second kind of counterexample applies to a whole class of important moral theories, which I call the class of *monistic extensional* moral theories. A moral theory is *monistic extensional* iff it entails that there is at least one deontic D and a property F such that:

1. $\Box\forall x[Fx \leftrightarrow Dx]$
2. $\Box\forall x[Dx \leftrightarrow E(Fx, Dx)]$
3. $\Box\forall x\forall G[E(Gx, Dx) \rightarrow F = G]$

In words, a monistic extensional moral theory is one which entails that there is *at least one* deontic property which is *uniquely explained* by a property with which it is *necessarily co-extensive*. Contractualism claims, for example, that wrongness is co-extensive with the contractualist property C , and it claims that wrong acts are wrong for the sole reason that they have property C . But, contractualism is not the only monistic extensional moral theory. Many other moral theories can be represented in the required form. Consider, for example, the following sets of properties:

\mathbb{C}_{rule} : {required by the best set(s) of rules, compatible with the best set(s) of rule, forbidden by the best set(s) of rules}

\mathbb{K} : {is an act the maxim of whose omission cannot be universalized, is an act whose maxim can be universalised, is an act whose maxim cannot be universalized}

\mathbb{V} : {would not be omitted by a virtuous agent, would be performed by a virtuous agent, would not be performed by a virtuous agent}

Within each set of properties, we might claim that the first property is co-extensive with obligation, the second with permissibility, and the third with wrongness. If we add that each of these properties uniquely explains the deontic property it is co-extensive with, then \mathbb{C}_{rule} corresponds to a

version of rule consequentialism, \mathbb{K} to a version of Kantianism, and \mathbb{V} to a version of first order virtue ethics. Any of these theories could be used in place of contractualism in presenting my second counterexample to the Biconditional Law Sufficiency Thesis. Kantians (as conceived here) want to claim the fact that a particular act is wrong is uniquely explained by the fact that its maxim cannot be universalized. If the Biconditional Sufficiency Thesis were true, this could not be correct. It would also be true that a wrong act is wrong because its maxim cannot be universalized and it is such that $2 + 2 = 4$.

Contractualism, Consequentialism, Kantianism and Virtue Ethics represent the four leading first order moral theories. My second kind of counterexample shows that the Biconditional Law Sufficiency Thesis is inconsistent with each of them. It therefore reduces to absurdity.

I have now raised two objections which I take to be fatal to the Biconditional Law Sufficiency Thesis, and hence also to sufficiency for deontic explanation of the weaker Strict Conditional analysis. Together, their failure marks the collapse of the analysis of normative explanation in terms of Modal Necessitation. Still, there remains the question of whether either of these analyses is *necessary* for normative explanation, which deserves some brief discussion.

4.3.2 Against the Strict Conditional Necessity

Thesis

In the previous section I argued that the Biconditional Law Analysis is insufficient for normative explanation. Might it nevertheless be necessary?

Let me suggest some reasons for doubt. Consider:

(4.3.9) **Strict Conditional Necessity Thesis**

$E(Fx, Dx)$ only if Fx and Dx and $\Box\forall z[Fz \rightarrow Dz]$.

According to 4.3.9, every normative explanation is backed by an exceptionless moral law of the form $\Box\forall z[Fz \rightarrow Dz]$. However, much recent work in both metaphysics, the philosophy of science and in ethics, has stressed the utility, and in some cases the indispensability, of a weaker conception of laws, often called *ceteris paribus* or *defeasible* laws.¹⁵ I will not argue the case for defeasible laws here. I note only that if laws of this weaker kind are possible, it would appear to undermine 4.3.9. 4.3.9 claims that every normative explanation must be backed by an exceptionless law. But, if defeasible laws are possible, that may not be true.

Consider now:

(4.3.10) **Biconditional Law Necessity Thesis**

$E(Fx, Dx)$ only if

1. Fx and Dx ; and

¹⁵Armstrong 1983, Ch. 10; Cartwright 1983, Ch. 2; Pietroski 1993; Pietroski and Rey 1995; Little 2000; Lance 2004; Lance and Little 2006b; Lance and Little 2006a; Bird 2007, § 3.3; Lance and Little 2007.

2. $\Box\forall z[Fz \leftrightarrow Dz]$; and
3. it is a law that $\Box\forall z[Fz \rightarrow Dz]$.

4.3.10 is stronger than 4.3.9, and vulnerable to further objections. In particular, 4.3.10 is inconsistent with a plausible moral pluralism in which a plurality of non-coextensive properties are each individually sufficient to make an act obligatory. Suppose, for example, that an act can be obligatory *either* because it is optimific *or* because there is no principle reasonably permitting its omission. Suppose, further, that every optimific act is one for which there is no principle reasonably permitting it, but that the reverse is not true, so that that every act which is obligatory on consequentialist grounds is also obligatory on contractualist grounds, but not vice versa. It follows that it is not the case that an act is obligatory iff it is optimific. So, there is no biconditional law connecting being optimific and being obligatory. This seems like a reasonable case of pluralism, and if it is, it directly contradicts condition (2) in 4.3.10. Of course, the possibility of pluralist ethical theories is a contested issue. Here I note only that 4.3.10 is inconsistent with some plausible versions of pluralism, and I add that it is debatable whether an account of normative explanation should be able to settle the question of the possibility of pluralism.

This completes my discussion of the first version of Metaphysicalism. I proceed now to discuss whether normative explanation can be analysed in terms of supervenience.

4.4 Supervenience Accounts of Normative Explanation

The second version of Metaphysicalism to consider is the thesis that normative explanation can be analysed in terms of *supervenience*. Supervenience is a relation between sets of properties in which the core idea is that of *covariance*¹⁶ or *dependent-variation*¹⁷. One of the lessons of the relatively recent work on supervenience is that there is a dense spectrum of such claims, varying in detail, generality and modal strength.¹⁸ Consequently, care is needed in formulating the supervenience relations to be discussed in this chapter. I operate with a standard notion of *strong, single domain* supervenience.¹⁹ Letting \mathbb{A} and \mathbb{B} be non-empty sets of properties, the strong version of supervenience may be defined as follows:

(4.4.1) \mathbb{A} *strongly supervenes on* \mathbb{B} =_{def.} for all possible worlds w and w' , and for all x and y , if x in w is \mathbb{B} -indiscernible from y in w' , then x in w is \mathbb{A} -indiscernible from y in w' .²⁰

To say that x and y are indiscernible with respect to the properties in a given set is just to say that for each property in that set, x has that property

¹⁶Kim 1993*b*, p. 140.

¹⁷McLaughlin 1995, p. 18.

¹⁸Contemporary discussion of weak, strong and global supervenience derives from Kim 1984, amended in Kim 1987. More sophisticated supervenience claims requiring multiple domains and co-ordinating relations are discussed in Kim 1993*c*. The wide variation in generality and modal strength of supervenience claims is discussed in, *inter alia*, Lewis 2001, esp. p. 14, McLaughlin 1995 and Charles 1992.

¹⁹A supervenience relation is *strong* iff its indistinguishability relation can relate entities in all metaphysically possible worlds. It is *single domain* iff it relates the distribution of sets of properties over a single, fixed domain of entities. (See Kim 1993*c*, p. 110)

²⁰Cf. Kim 1984; McLaughlin 1995.

iff y also has that property. I discuss some further technicalities relating to this definition below, when needed.

The idea that normative explanation might be analysed in terms of supervenience has proved especially attractive. Jaegwon Kim has suggested that ‘the belief that the moral supervenes on the non-moral may have shaped some of the major assumptions and tasks of moral philosophy’, including ‘the belief that there are such things as “good-making” or “right-making characteristics”’.²¹ Mackie, despite his scepticism regarding normative explanation, wrote that objective deontic properties, if they exist at all, ‘must be somehow “consequential” or “supervenient”’²² on the properties that explained them. Some have claimed outright that supervenience answers Mackie’s sceptical question concerning normative explanation. Walter Sinnott Armstrong writes:

Besides, it’s not that hard to specify the relation between moral facts and nonmoral facts. That relation is supervenience.²³

Confidence in a supervenience analysis of normative explanation may also explain the motivation behind various attempts to define first order moral theories using supervenience. For, as I argued in § 1.3.1, first order moral theories are simply comprehensive accounts of what explains why acts are right and wrong. One example of a supervenience definition of a moral theory is to be found in the work of G.E. Moore. In § 4.3 I discussed Moore’s two-part definition of direct act utilitarianism in his *Ethics*. Later

²¹Kim 1984, p. 176.

²²Mackie 1977, p. 41.

²³Sinnott Armstrong 2006, p. 50.

in that work Moore supplements his official definition of utilitarianism with a further principle he takes to be central to the theory:

Let us suppose [...] that we have an action X, which is right, and whose total effects are A; and let us suppose that the total effects of all the possible alternative actions would have been respectively B, C, D and E. The precise principle with which we are now concerned may then be stated as follows. Our theory [utilitarianism] implies, namely, that any action Y which resembled X in both the two respects (1) that its total effects were precisely similar to A and (2) that the total effects of all the possible alternatives were precisely similar to B, C, D and E, would necessarily also be right, if X was right, and would necessarily also be wrong, if X was wrong.²⁴

Bykvist²⁵ explicitly develops Moore's principle as a supervenience definition of consequentialism, and it is natural to wonder whether, so understood, it might supplement Moore's definition in a way that would capture the elusive explanatory element that he recognised was missing from the definition given in terms of biconditional laws. Supervenience definitions of consequentialism, or a distinctive element in its core, have also been developed by Vallentyne²⁶ and Broome²⁷. However, all three authors remain sceptical about how far such definitions really do capture the idea of normative explanation.

Part of the enthusiasm for supervenience accounts of normative explanation stems from the widely held belief that explanation in general has something to do with supervenience. But, there are also motivations particular to the moral case. First, there is near universal agreement among

²⁴Moore 1912, p. 84.

²⁵Bykvist 2003.

²⁶Vallentyne 2006.

²⁷Broome 2004*b*, p. 32.

philosophers that moral properties supervene on nonmoral properties. Call this claim *Moral Supervenience*.²⁸ Second, there is near universal agreement that Moral Supervenience is an *a priori* truth, perhaps because it is a conceptual truth. Nick Zangwill aptly summarizes how these claims have bolstered the motivation for a supervenience analysis of normative explanation:

A plausible thought is that we need a *necessity* to make sense of the “because”: moral and natural properties are necessarily connected. What is called moral-natural “supervenience” is the usual way of articulating precisely what the necessary connection is.²⁹

The moral of this brief introduction is that the motivations for a supervenience analysis are complex. So before I proceed, it will be wise to clarify their significance. I recognise three motivations for a supervenience analysis of normative explanation. First, there is Moral Supervenience, the claim that moral properties strongly supervene on nonmoral properties. Moral Supervenience is a valid consideration and one that I will take into account in my discussion, with one refinement that I mention below. Second, it is widely accepted that Moral Supervenience is knowable *a priori*. This claim is not relevant to my discussion. The claim of *a priori* tells us something about how we can *know* that the moral supervenes on the nonmoral, but nothing about whether it can be used to analyse normative explanation.

²⁸Some prefer to claim that moral properties *globally supervene* on nonmoral properties, formulated as the claim that: for all possible worlds *w* and *w'*, if *w* and *w'* are descriptively indistinguishable then they are morally indistinguishable. Global supervenience is, in fact, weaker than strong supervenience (cf. Kim 1987), but for my purposes, the difference between them is negligible.

²⁹Zangwill 2008.

Third, there is the idea that Moral Supervenience is a conceptual truth. This claim is not relevant to the present analysis of normative explanation in terms of *metaphysical* necessity, but it is relevant to Conceptualism, the thesis that normative explanation might be analysed in terms of *conceptual* necessity. So, I postpone discussion of the alleged conceptual truth of Moral Supervenience until § 4.5. For now, I focus on the question of whether normative explanation can be analysed in terms of a strong supervenience relation, but only insofar as this relation captures some kind of *metaphysical* necessity.

Moral Supervenience claims that an entire set of moral properties, including properties of overall and non-overall normative status, evaluative properties and areteic properties, supervene on nonmoral properties. I acknowledged above that this claim is directly relevant to a supervenience analysis of normative explanation, but for that purpose, Moral Supervenience is broader than is really needed. For present purposes, all that matters is that the overall deontic properties of obligation, wrongness and permissibility supervene on nonmoral properties. The supervenience of the deontic properties on nonmoral properties follows from Moral Supervenience because supervenience has the notable feature that, if set \mathbb{A} supervenes on set \mathbb{B} , then *every subset* of \mathbb{A} also supervenes on \mathbb{B} . So, if moral properties supervene on nonmoral properties, then the set of overall deontic properties, which is a subset of the moral properties, also supervenes on nonmoral properties. Call the claim that the overall deontic properties supervene on nonmoral properties *Deontic Supervenience*. In what follows, I focus on this more restricted supervenience claim.

4.4.1 Formulating the Supervenience Analysis

Before the supervenience analysis of normative explanation can be assessed, its correct form must be found. My schema for this analysis will be:

(4.4.2) **Supervenience Analysis Schema**³⁰

$E(\{F_1x, \dots F_ix\}, Dx) \leftrightarrow$

1. F_1x and $\dots F_ix$; and
2. Dx ; and
3. $S(\{F_1, \dots F_i\}, D)$

4.4.2 expresses the idea that an act has deontic property D because it has one or more properties $F_1, \dots F_i$ iff the act has properties $F_1, \dots F_i$, it also has property D , and finally, there is a between them a relation, S , which is to be cashed out in terms of supervenience. This schema permits a range of options.

Most discussions of supervenience follow Kim in assuming that the operations of negation, disjunction and conjunction (the *Boolean operations*) are valid methods of forming properties, so that if F and G are properties, then so, too, are not- F and not- G , F -and- G , and F -or- G . It is also assumed that supervenience relates sets of properties closed under these operations. Given these further assumptions, my earlier definition of supervenience in 4.4.1 is necessarily equivalent to the alternative *Modal Operator* formulation of supervenience:

³⁰I am grateful to my supervisor for help in developing this schema.

(4.4.3) Modal Operator Supervenience

\mathbb{A} strongly supervenes on \mathbb{B} iff necessarily, for each x and each property A in \mathbb{A} , if x has A , then there is a property B in \mathbb{B} such that x has B , and necessarily, if any y has B , it has A .³¹

I will begin by assuming that the Boolean operations are valid. So I will employ the Modal Operator definition of supervenience in formulating my analysis. I note, however, that the validity of the Boolean operations has been criticized on several grounds.³² I will discuss later whether it would be worth rejecting them.

The assumption that the Boolean property-formation operations are valid has important implications for the supervenience analysis and raises a number of questions. Suppose that the deontic properties supervene on a set of properties, \mathbb{F} . On the assumption that \mathbb{F} is closed under the Boolean operations, every act must have a number of properties from \mathbb{F} . First, for each atomic (i.e. non-Boolean) property in \mathbb{F} , the act must have either it or its negation. Second, the act must have exactly one \mathbb{F} -*maximal* property, where the \mathbb{F} -maximal properties are the strongest consistent properties constructible in \mathbb{F} . The \mathbb{F} -maximal properties are mutually exclusive and every object must have exactly one of them.³³ Third, for any given deontic property, D , in the supervening set, we can also speak of its \mathbb{F} -*minimal* base, which is the weakest property constructible in \mathbb{F} such that necessarily, if any

³¹Cf. Kim 1984, p. 165.

³²Van Cleve 1990; Armstrong 1978; McLaughlin 1995, p. 21.

³³Kim 1984, p. 158.

act has that property, then it has D .³⁴ This raises an important question. Given any act with a particular deontic status, such as being wrong, which \mathbb{F} -properties should we take to explain why it is wrong? There are three salient possibilities. We could say that the act is wrong because:

1. it has precisely those atomic properties in \mathbb{F} which it has; or
2. it has its \mathbb{F} -maximal property; or
3. it has the \mathbb{F} -minimal property for wrongness.

To simplify my exposition, I will begin by taking the \mathbb{F} -maximal property to be relevant explainer. But, I will show later that nothing, in fact, rests on this choice.

A second issue to resolve in developing a supervenience analysis of normative explanation is whether the relevant supervenience relation must be asymmetric. I mentioned above the widely accepted claim of Moral Supervenience, according to which moral properties supervene on nonmoral properties. While supervenience is itself a non-symmetric relation (it is neither symmetric nor asymmetric), the supervenience of the moral on the nonmoral is usually understood to include the idea that the nonmoral properties do not also supervene on the moral properties. I will start with the simplifying assumption that the relevant supervenience component need not be asymmetric. An amendment involving asymmetric supervenience will be considered in due course.

³⁴Kim 1984, p. 158. There can be only one \mathbb{F} -minimal property for any given deontic property. For suppose two such properties, F_1 and F_2 each necessitated deontic property D . Then their disjunction, F_1 -or- F_2 would also necessitate D . But the disjunction of F_1 and F_2 is logically weaker than either property alone. (Cf. Wedgwood 2007, p. 151)

I can now propose a first supervenience analysis of normative explanation. Let \mathbb{N} be the set of descriptive properties, and let \mathbb{D} be the set of overall deontic properties.

(4.4.4) **Supervenience Analysis**

For all properties F , acts x and deontic properties D : $E(Fx, Dx) \leftrightarrow$

1. Fx and Dx ; and
2. $\Box\forall y[Fy \rightarrow Dy]$; and
3. F is a maximal property of \mathbb{N} ; and
4. the set of deontic properties \mathbb{D} strongly supervenes on \mathbb{N}

This analysis may be understood as follows. Condition 1 is the usual requirement of factiveness. Condition 3 is the requirement that the explaining property be a *maximal* descriptive property. Condition 4 is the requirement that deontic properties strongly supervene on descriptive properties. Strictly speaking, condition 2 in this analysis is redundant, because it is entailed by conditions 1, 3 and 4. But, its addition brings greater clarity at no cost.

Let the *Supervenience Sufficiency Thesis* and *Supervenience Necessity Thesis* be the claims, respectively, that the Supervenience Analysis in 4.4.4 is sufficient and necessary for normative explanation. My major aim is to argue that the Supervenience Sufficiency Thesis is false and unsalvageable. I will then argue that doubts can be raised about the Supervenience Necessity Thesis. I should clarify, perhaps, that none of what I say in this chapter

is intended to deny the claims of Moral or Deontic Supervenience. On the contrary, I assume that they are true throughout this thesis.

4.4.2 Against the Supervenience Sufficiency Thesis

Let the *Supervenience Sufficiency Thesis* be the claim the Supervenience Analysis in 4.4.4 is sufficient for normative explanation, i.e.:

(4.4.5) Supervenience Analysis

For all properties F , acts x and deontic properties D , if

1. Fx and Dx ; and
2. $\Box\forall y[Fy \rightarrow Dy]$; and
3. F is a maximal property of \mathbb{N} ; and
4. the set of deontic properties \mathbb{D} strongly supervenes on the \mathbb{N}

then $E(Fx, Dx)$.

I assume that it is uncontroversially true that the deontic properties supervene on descriptive properties. Given that assumption, we can take the *conditional* of Deontic Supervenience and 4.4.5 as follows:

(4.4.6) If the set of deontic properties \mathbb{D} supervenes on set \mathbb{N} then, if for all properties F , acts x and deontic properties D :

1. Fx and Dx ; and
2. $\Box\forall y[Fy \rightarrow Dy]$; and

3. F is a maximal property of \mathbb{N} ; and
4. the set of deontic properties, \mathbb{D} , supervenes on \mathbb{N}

then $E(Fx, Dx)$.

It is now easily seen that condition 4 of 4.4.6 is redundant, since it is equivalent to the claim of Deontic Supervenience stated in the antecedent. Hence 4.4.6 entails:

(4.4.7) If the set of deontic properties \mathbb{D} supervenes on set \mathbb{N} then, *if* for all properties F , acts x and deontic properties D :

1. Fx and Dx ; and
2. $\Box\forall y[Fy \rightarrow Dy]$; and
3. F is a maximal property of \mathbb{N} ; and

then $E(Fx, Dx)$.

Granting, again, that the deontic properties supervene on descriptive properties, we can detach the consequent of 4.4.7, to yield:

(4.4.8) For all properties F , acts x and deontic properties D :

1. Fx and Dx ; and
2. $\Box\forall y[Fy \rightarrow Dy]$; and
3. F is a maximal property of \mathbb{N} ; and

then $E(Fx, Dx)$.

4.4.8 shows that, despite its initial complexity, the Supervenience Sufficiency Thesis reduces to a mild strengthening of the Strict Conditional Analysis of normative explanation, 4.3.1, considered above in § 4.3. 4.4.8 shares conditions 1 and 2 with the Strict Conditional Sufficiency Thesis. It merely adds the further claim that property F is maximal in set N . But, this yields no improvement. We can verify that 4.4.8 is vulnerable to the same objections that were fatal to the sufficiency of the Strict Conditional analysis as follows.

Consider the set of properties:

\mathbb{G} : {God commands x , God forbids x }

Let \mathbb{G}_{Clos} be the smallest superset of \mathbb{G} that is closed under the Boolean operations. It is conceptually possible that the properties in \mathbb{G} are necessarily co-extensive with being obligatory, and wrong, respectively, and that the property of being permissible is necessarily co-extensive with lacking the property of being forbidden by God. If this is so, it follows that the deontic properties supervene on \mathbb{G}_{Clos} . Now take the maximal \mathbb{G}_{Clos} -property that can be constructed from the property that God commands x and the negation of the property that God forbids x . Then, by 4.4.4, this property would explain why obligatory acts are obligatory. But, as I have objected before, we need not accept that divine command explains why acts are right or wrong. So, we have a first counterexample to the Supervenience Sufficiency Thesis.

For a second example, consider again the set of properties (introduced in § 4.3.1):

\mathbb{V} : {would not be omitted by a virtuous agent, would be performed by a virtuous agent, would not be performed by a virtuous agent}

Let \mathbb{V}_{Clos} be the smallest Boolean closure of this set. Second Order Virtue Theory, of the sort I described in 1.3.2, entails that the properties in \mathbb{V} are necessary and sufficient for obligation, permissibility and wrongness respectively. So, the deontic properties supervene on \mathbb{V}_{Clos} . But, Second Order Virtue Theory denies that having these properties explains why acts have the corresponding deontic properties. So we have a second counterexample to the sufficiency of the Supervenience Analysis. In summary, 4.4.4 fails to be sufficient for the same reason as did our previous analyses in terms of strict conditionals. I will now consider three replies.

I noted above that there are many different way of formulating a supervenience analysis of normative explanation. However, the argument that I have just given can be extended to show that two major alternative formulations will also fail to be sufficient for normative explanation. The important point to note in extending the above argument is that it need not be accepted that *any* of the properties in set G , nor in its Boolean closure \mathbb{G}_{Clos} , normatively explains a deontic property. This allows me to extend my argument to cover the following variations of the supervenience analysis.

First, I noted above that, under the assumption of the validity of the Boolean operations, we might take an act's deontic status to be explained by either the atomic properties it has from the supervenience base, or by its maximal property from that base, or by the weakest property in the

supervenience base which necessarily implies the deontic status in question. The fact that *none* of the properties on \mathbb{G} , and *none* of the properties in its Boolean closure \mathbb{G}_{Clos} , need be taken to explain an act's deontic status, shows that there are counterexamples to sufficiency of the supervenience analysis no matter which of these options is taken.

Second, in developing the Supervenience Analysis in 4.4.4, I assumed the validity of the Boolean property formation operations, and I assumed that supervenience related sets of properties closed under these operations. So, it might be objected that the supervenience analysis would fare better without them. I deny this claim. If we relax these assumptions, then the deontic properties supervene directly on property set \mathbb{G} . But, it need not be accepted that those properties explain why acts are right or wrong.

The third variation of the supervenience analysis must be handled in a different way. In developing the Supervenience Analysis in 4.4.4, I assumed that the supervenience relation need not be asymmetric. So, it might be objected that the supervenience analysis would fare better if it required asymmetric supervenience instead. This variation looks initially promising because the deontic properties *symmetrically* supervene on the divine command properties in \mathbb{G}_{Clos} . So, by adding the requirement of asymmetry, the objections deriving from the divine command theory can be avoided. The problem with this response, however, is that many plausible moral theories imply that the deontic properties *symmetrically* supervene on the properties that explain them.³⁵ Consider, for example, this set of properties:

³⁵Similar objections to the asymmetry condition are made by Depaul 1987, p. 434; Bykvist 2003, p. 10. I developed my objection independently from Depaul, and there are differences between his and my own. Depaul considers a supervenience analysis whereby

\mathbb{C}_{act} { x is uniquely optimific, x is optimific, x is sub-optimific}

A version of maximising consequentialism might be defined so that (1) an act is obligatory iff it is uniquely optimific, permissible iff it is optimific, and wrong iff it is sub-optimific, and (2) each deontic property is explained by the property with which it is co-extensive. However, this definition entails that the deontic properties *symmetrically* supervene on set \mathbb{C}_{act} . So, if we add to our supervenience analysis the requirement that deontic properties must *asymmetrically* supervene on the properties that explain why acts are right or wrong, we will thereby rule out a consequentialist account of what explains why acts are right or wrong. Furthermore, this problem is by no means limited to act consequentialism. Consider, again, the sets of properties:

\mathbb{C}_{rule} : {required by the best set(s) of rules, compatible with the best set(s) of rule, forbidden by the best set(s) of rules}

\mathbb{K} : {is an act the maxim of whose omission cannot be universalized, is an act whose maxim can be universalised, is an act whose maxim cannot be universalized}

\mathbb{V} : {would not be omitted by a virtuous agent, would be performed by a virtuous agent, would not be performed by a virtuous agent}

each individual deontic property is taken to be explained by a single property on which it supervenes, e.g. the property of being optimific. My supervenience analysis relies instead on a supervenience relation between the set of deontic properties and a second set of properties. I also go beyond Depaul in considering the effect of denying the validity of the Boolean operations. In spirit, though, our objections are similar.

We can define versions of Rule Consequentialism, Kantianism and Virtue Ethics according to which acts are obligatory iff they have the first listed property in the relevant set, permissible iff they have the second listed property, and wrong iff they have the third. In each case, the deontic properties will symmetrically supervene on the set of properties in question. So, the supervenience analysis can require asymmetric supervenience only at the cost of ruling out most of the leading, first order moral theories. This reduces the asymmetric supervenience analysis to absurdity.

This completes my argument against the Supervenience Sufficiency Thesis. I have considered some of the principal variations of the supervenience analysis and I have argued that each of them either fails to be sufficient or reduces to absurdity. I will now briefly discuss whether the Supervenience analysis is *necessary* for normative explanation.

4.4.3 Against the Supervenience Necessity Thesis

Let the *Supervenience Necessity Thesis* be the claim that the Supervenience Analysis in 4.4.4 is necessary for normative explanation, i.e.

(4.4.9) For all properties F , acts x and deontic properties D , if $E(Fx, Dx)$

then

1. Fx and Dx ; and
2. $\Box\forall y[Fy \rightarrow Dy]$; and
3. F is a maximal property of \mathbb{N} ; and
4. the set of deontic properties, \mathbb{D} , strongly supervenes on \mathbb{N}

I argue that doubts can be raised about 4.4.9. Consider conditions 1–4 of the analysis. I have argued that normative explanation is factive, so condition 1 is clearly correct. Furthermore, it is assumed throughout this thesis that the deontic properties supervene on nonmoral properties. So condition 4 is clearly correct. So, if there are difficulties, they concern conditions 2–3.

Consider condition 3. With this condition in place, 4.4.9 claims that an act's being F explains why it is e.g. obligatory only if F is the act's maximal natural property. It is well recognised, however, that maximal properties will include irrelevant conjuncts. For example, they will include how many hairs there were on the agent's head, whether the act was performed in the vicinity of some chewing gum, and so on. So, we should reject the condition that explaining properties must be maximal.

I noted above that the usual alternative is to take F to be the logically weakest natural property that necessarily implies the deontic property whose instantiation is to be explained. But, this, too, will bring difficulties. Consider a simple form of utilitarianism according to which obligatory acts are obligatory because they are uniquely felicific, and according to which an act is obligatory iff it is uniquely felicific. However, let us assume that God commands only uniquely felicific acts, and consider the complex Boolean property of being uniquely-felicific-or-divinely-commanded. This property is logically weaker than the property of being uniquely-felicific. So this version of the Supervenience Necessity Thesis is inconsistent with utilitarianism. Utilitarians, so this theory would imply, should instead claim that acts are right because they are uniquely-felicific-or-divinely commanded. Once again, this problem applies not only to utilitarianism, but to any

moral theory with a similar structure.

One way to avoid this difficulty would be as follows. To cope with the kind of difficulties I have raised above, Ralph Wedgwood has argued that we should focus instead on the minimal, *non-disjunctive* property in any supervenience base.³⁶ This cannot be quite right because there may be more than one such property. For example, the property of being divinely commanded may be one such minimal base, the property of being felicitic another. To avoid this difficulty, we might stipulate that *F* should be *among* the minimal, non-disjunctive supervenience bases for the relevant deontic property *D*, however, the property of being divinely commanded will always be *among* the minimal, non-disjunctive supervenience bases for obligatory acts, so there seem to be ineliminable difficulties with condition 3.

The final difficulty to raise with regard to 4.4.9 is condition 2. This requires that explaining properties necessarily imply the properties they explain. As I noted in my discussion of the Strict Conditional Biconditional Law Necessity Thesis 4.3.2, the idea that explaining properties necessarily imply the properties they explain has been questioned on the basis that moral laws might take a defeasible form. For example, the fact that I promised to help might explain why I have an obligation to help only under the condition that I promised *freely*. If my promise is coerced, it may not necessarily give rise to any obligation.

In this section, I have considered two versions of Metaphysicalism, one in terms of Metaphysical Necessitation, the second in terms of Supervenience. I have argued that neither version is sufficient for normative explanation,

³⁶Wedgwood 2007, p. 151.

and I have suggested some reasons to doubt that either version is necessary. I argued that the Supervenience Analysis failed for much the same reasons as did the Metaphysical Necessitation Analysis, because it reduces to only a mild strengthening of that claim. The failure of each of these versions of Metaphysicalism strongly suggests that no reduction of normative explanation to metaphysical necessity is possible. I complete my chapter with a brief discussion of whether normative explanation can be analysed in terms of *conceptual necessity*.

4.5 Conceptualism

My aim in this section is to provide some brief arguments against one of the remaining possibilities for a reduction of normative explanation to non-normative terms. *Conceptualism*, the view to be assessed here, claims that normative explanation can be reduced to conceptual necessity. Here, I briefly criticize the two most prominent possibilities, which I call *Direct* and *Indirect* Conceptualism.

4.5.1 Direct Conceptualism

Direct Conceptual Reduction is the thesis that normative explanation can be reduced to a direct conceptual connection between the explanans and explanandum. This is best illustrated by example, so I will discuss the version of *Analytic Consequentialism* proposed by G.E. Moore in *Principia Ethica*³⁷. Moore writes:

³⁷Moore 1993.

In short, to assert that a certain line of conduct is, at a given time, absolutely right or obligatory, is obviously to assert that more good or less evil will exist in the world, if it be adopted than if anything else be done instead.³⁸

I understand Moore to be claiming that the concept ‘*x* is obligatory’ is identical with the concept ‘*x* is optimific’. The claim is not merely that these two concepts have the same *extension*, for that could be true even if they were not analytically connected. Moore’s claim is rather that these two concepts have identical *intension*, so they are identical *concepts* (and hence they pick out identical *properties*).³⁹ The connection between being obligatory and being optimific is, therefore, one of conceptual necessity.

Moore’s claim has been criticised on several grounds. Moore himself later rejected Analytic Consequentialism on the grounds that it failed the test posed by the *Open Question* argument, an objection also made by Ross.⁴⁰ The Open Question argument is, however, notoriously difficult to understand, and though many have thought there was something important in it, no consensus has been reached in identifying what it is.

Preoccupation with the Open Question argument may have insulated Analytic Consequentialism from more obvious objections. Consider the claim that obligatory acts are obligatory *because* they are optimific. It is trivial logical truth that this claim must be either true or false, but Analytic Consequentialism faces difficulty either way. It is natural for a consequentialist to accept the first horn of the dilemma and claim that obligatory acts are obligatory because they are optimific. But, if Analytic Consequential-

³⁸Moore 1993, § 17, p. 17.

³⁹Moore also claimed that the identity of these concepts was ‘obvious’, but I ignore this additional claim.

⁴⁰Ross 2002, p. 16.

ism is true, the property of being optimific is *identical* with the property of being obligatory. So, it follows that obligatory acts are obligatory *because* they are obligatory. In other words, Analytic Consequentialism and first order consequentialism together entail that for every act x , if x is obligatory, then the fact that x is obligatory normatively explains why x is obligatory. But, this claim should be denied on the grounds of pernicious reflexivity. It is never true that an act is obligatory because it is obligatory. The perhaps surprising result is, then, is that Analytic Consequentialism is inconsistent with first order consequentialism. If ‘ x is obligatory’ just means ‘ x is optimific’, then the fact that x is optimific can provide no normative explanation of why x is obligatory. The Analytic Consequentialist must therefore deny first order consequentialism.

If the Analytic Consequentialist denies first order consequentialism, two possibilities remain. First, it might be denied that *anything* normatively explains why obligatory acts are obligatory. To take this option is to deny the claim of *Explanatory Moral Dependence* (ExtDep 1) that I defended in § 2.4.1. However, many take Explanatory Moral Dependence itself to be a conceptual truth of the most fundamental sort, so this would be a very high cost to pay. The only remaining alternative for the Analytic Consequentialist is to claim that something *other* than being optimific normatively explains why obligatory acts are obligatory. However, this last alternative brings with it two significant problems of its own. First, Analytic Consequentialism entails that an act is obligatory iff it is optimific, so Analytic Consequentialism is compatible with only the very limited range of first order moral theories that still respect that constraint. Second, and

more problematically, the combination of Analytic Consequentialism with a non-consequentialist first order theory looks unstable. Consider, for example, the combination of Analytic Consequentialism with first order divine command theory, according to which obligatory acts are obligatory because they are commanded by God. If one accepts first order divine command theory, it is hard to see any plausibility in the claim the concept of *being obligatory* is identical to the conception of *being optimific*. Analytic Consequentialism draws whatever plausibility it has from the *first order* claim that obligatory acts are obligatory *because* they are optimific. Combining Analytic Consequentialism with a different first order moral theory simply undermines that initial plausibility.

The objections I have just outlined to Analytic Consequentialism are symptomatic of all Analytic Conceptual Reductions of normative explanation. An Analytic Conceptual Reduction claims that the concept of *being obligatory* is identical to the concept of *being F*, for some value of *F*. If we accept that normative explanation is irreflexive, then for every *x*, it is false that *x* is obligatory *because x is F*. The Analytic Reduction of Obligation to *F*ness is thus inconsistent with any first order normative connection between them. So, the defender of the analytic view must either pay the very high cost of denying that any normative explanations exist, or else defend an alternative first order theory. But, if obligation has a first order connection with some property *other* than *F*, there is no reason to suppose that obligation has a conceptual connection to *F*ness. For these reasons, I deny that normative explanation is susceptible to Analytic Conceptual Reduction.⁴¹

⁴¹I am grateful to my supervisor for help in developing this line of objection.

4.5.2 Indirect Conceptualism

The alternative to Direct Conceptual Reduction is *Indirect Conceptual Reduction*. Direct Conceptual Reduction of normative explanation was characterized by a direct conceptual connection between the explanans and explanandum of the normative explanation. The alternative strategy is to claim that it is a *conceptual* truth that there is some *non-conceptual* connection between explanans and explanandum, and that this *non-conceptual* connection is sufficient (and perhaps also necessary) for normative explanation. In this way, it might be claimed that normative explanation is underwritten by an *indirect* conceptual connection.

The most prominent forms of Indirect Conceptual Reductionist argument that one finds in the literature rest on considerations of moral-nonmoral supervenience. A representative version of this argument would run:

1. It is a conceptual truth that deontic properties strongly supervene on some descriptive properties.
2. It is metaphysically necessary (and perhaps also a conceptual truth) that if deontic properties strongly supervene on a set of properties \mathbb{F} , then an act's \mathbb{F} -properties (or some subset of them) normatively explain why it has its deontic properties.

Therefore:

- C. An act's descriptive properties (or some subset of them) normatively explain why it has its deontic properties.

Popular though this argument is, it should be rejected. I argued at length in § 4.4 that the second premiss of this argument is false. Supervenience, though perhaps necessary for normative explanation, is insufficient for it. So this argument, though valid, is unsound.

Some may find my argument against Indirect Conceptual Reduction too quick, so I will consider a further example of indirect conceptual reduction at greater length. In *From metaphysics to ethics : a defence of conceptual analysis*⁴², Frank Jackson presents an argument that I classify as a version of Indirect Conceptualism. Jackson's argument proceeds in several stages, and my objections concern only the final stage. The salient points of the first stage of Jackson's argument are as follows.

1. (Premiss) Moral properties supervene on descriptive properties.
2. (Premiss) Necessarily, if \mathbb{A} -properties supervene on \mathbb{B} -properties, then for every \mathbb{A} -property there is some \mathbb{B} -property that is necessarily co-extensive with it.⁴³
3. (From 1, 2) For every moral property there is some descriptive property that is necessarily co-extensive with it.
4. (Premiss) Necessarily co-extensive properties are identical.

Therefore:

⁴²Jackson 1998.

⁴³Jackson's presentation of this argument is premised on the *global* supervenience of the moral on the descriptive (see Jackson 1998, p. 119). However, as Williamson (2001) argues, the global supervenience claim is too weak to support the second premiss's claim that every supervening property is necessarily equivalent with some base property. What is needed instead is that moral properties *strongly* supervene on descriptive properties. I am happy to grant that assumption here.

C. Every moral property is identical with a descriptive property.

If the premisses of this argument are taken to hold of conceptual necessity, then its conclusion is also a conceptual necessity.

Jackson's argument can be criticized on several grounds.⁴⁴ In particular, one can question Jackson's highly controversial premiss that necessarily co-extensive properties are identical.⁴⁵ However, rather than questioning the soundness of Jackson's argument, I will criticize the significance of his conclusion. Let us grant for the sake of argument that every moral property is identical with a descriptive property, and ask what follows.

Having established that moral properties are identical with descriptive properties, Jackson proceeds to the second stage of his argument, in which he proposes to identify moral properties and descriptive properties using the method of Ramsey sentences defended by David Lewis.⁴⁶ To run this argument, we are required to imagine there is such a thing as 'mature folk morality', which is the result of extensive improvements to existing folk morality. Let M be mature folk morality, written as a long conjunction, with each moral predicate written in property-name form e.g. 'Act a has rightness' instead of 'Act a is right'. Then, we replace each distinct moral term in M with a distinct variable, giving $M(x_1, \dots, x_i)$. From $M(x_1, \dots, x_i)$, we form the following Ramsey sentence, which says that M has a unique realization:

$$(4.5.1) \quad \exists x_1 \dots \exists x_i \forall y_1 \dots \forall y_i [M(y_1, \dots, y_i) \leftrightarrow [x_1 = y_1 \wedge \dots \wedge x_i = y_i]]$$

⁴⁴Cf. Roojen 1996.

⁴⁵Critics of this premiss include Shafer-Landau 2003, Ch. 4; Majors 2005. Defenders include Streumer 2008.

⁴⁶Jackson 1998, pp. 140–142; Cf. Lewis 1970.

Jackson argues that M and 4.5.1 ‘say the same thing’, if, but only if, it is part of mature folk theory that it has a unique realization.⁴⁷ That mature folk morality has a unique realization is doubtless a further controversial assumption of Jackson’s argument, but, once again, I shall overlook it in order to reach the heart of Jackson’s position.

Granting the equivalence between M and 4.5.1, Jackson notes that we can now say what it is for an act a to be e.g. right as follows:

$$(4.5.2) \text{ [Action } a \text{ is right} \leftrightarrow \exists x_1 \dots \exists x_i [a \text{ has } x_1 \wedge \\ [\forall y_1 \dots \forall y_i [M(y_1, \dots, y_i) \leftrightarrow [x_1 = y_1 \wedge \dots x_i = y_i]]]]]$$

where ‘ x_1 ’ replaced ‘rightness’ in M . The section of Jackson’s argument I will criticize reads as follows:

This in itself does not tell us what rightness, the property, is, and the same goes for goodness, etc. It is a story about truth-conditions, but does not tell us about the metaphysics of rightness. In particular, it leaves open two possibilities: that rightness is the (first-order) descriptive property that plays the rightness role, the realizer property as it is called in the corresponding debate in the philosophy of mind, or that it is the second-order property of having the property that plays the rightness role, the role property as it is called in the philosophy of mind. However, there seems to be a clear reason for favouring the first view. We want rightness to be what makes an action right, not in the causal sense but in the sense of what ought to be aimed at. Now what we should aim at is not doing what is right *qua* what is right. I should rescue someone from a fire because if I don’t they will die, not because that is the right thing to do. [...] what ought to motivate us, and what we should value and pursue, is not the moral status of our actions *per se*, but the goods that confer moral status. But, from the perspective of

⁴⁷Jackson 1998, p. 140 and footnote 3.

moral functionalism, the choice between role property and realizer property is the choice between the moral property *per se* and what makes something right in the sense of being the rightness part of the of the best solution to the equations of mature folk morality [...].⁴⁸

I take Jackson to be claiming, on the basis of this argument, to identify what *makes*, or in other words, *explains why* right acts are right. So let us consider this argument more closely. The salient points seem to be these:

1. (Premiss) Rightness is the unique property that one ought to aim at.
2. (Premiss) For all properties x , x is the unique property that normatively explains why right acts are right iff property x is the unique property that one ought to aim at.
3. (From 1, 2) Rightness is the unique property that normatively explains why right acts are right.
4. (Premiss) Either (a) rightness is the unique property that is referred to by 'rightness' in the unique realization of M ; or (b) rightness is the (second-order) property of having the unique property that is referred to by 'rightness' in the unique realization of M .
5. (Premiss) The unique property that one ought to aim at is the unique property that is referred to by 'rightness' in the unique realization of M .

Therefore:

⁴⁸Jackson 1998, p. 141.

6. The unique property that is referred to by ‘rightness’ in the unique realization of *M* is the unique property that normatively explains why right acts are right.

I will argue that Jackson’s premisses are ambiguous, and that disambiguating them shows his argument to be unsound.

I start with premisses 1 and 2. These premisses are my rendering of Jackson’s claim that ‘we want rightness to be what makes an action right [...] in the sense of what ought to be aimed at’. Consider first premiss 1. This can be understood either a substantive claim about the property of *rightness*, using that word with its established meaning, or, as a new technical definition of the meaning of ‘rightness’. It is clear, however, it must be taken as a substantive claim about *rightness* with its current meaning, otherwise the attempt to identify rightness by applying Lewis’ method of definition to folk morality is misguided.

Taking premiss 1 to be substantive, we can now assess whether it is true. To answer that question, we need to understand what it means to say that a property ‘ought to be aimed at’. It is clear that, if Jackson’s argument is to be valid, the meaning of being a property that ‘one ought to aim at’ must remain constant between premisses 1 and 2. So let us consider premiss 2.

Premiss 2 proposes an equivalence between the property that normatively explains why right acts are right, and the property that one ought to aim at. However, it is ambiguous which property is the property *one ought to aim at*. What one ought to aim at might be:

- (i) something that one ought consciously to try to achieve, i.e. something

that ought to motivate us; or

(ii) something that normatively explains why one ought to do something;

or

(iii) something that explains why we ought to do something in some non-normative, perhaps formal mode.

The distinction between (i) and (ii) is the standard distinction through which self-effacing moral theories such as consequentialism are defined, and it standard to accept that what ought to motivate us and what normatively explains why we ought to act can come apart. As Sidgwick writes:

[...] it is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim: and if experience shows that the general happiness will be more satisfactorily attained if men frequently act from other motives than pure universal philanthropy, it is obvious that these other motives are reasonably to be preferred on Utilitarian principles.⁴⁹

The claim that morality (and rationality) has a *formal* aim is made by Parfit:

We can describe all [practical] theories by saying what they tell us to try to achieve. According to all moral theories, we ought to try to act morally. According to all theories about rationality, we ought to try to act rationally. Call these our *formal* aims.⁵⁰

Given Parfit's defence of Sidgwick's distinction between what ought to motivate us and what normatively explains why acts are right, it seems clear

⁴⁹Sidgwick 1962, p. 413.

⁵⁰Parfit 1987, p. 3.

that Parfit's *formal* aim is not intended to be what we ought *consciously* to aim at. Nor, however, is Parfit's *formal* aim intended to capture the idea of what normatively explains why acts are right or wrong. For that, Parfit reserves the term *substantive* aim. We might, then, take Parfit's idea of a formal aim to capture some other mode of explanation of why acts are right or wrong — a *formal* mode of explanation, perhaps.

Given that premiss 2 is ambiguous, how should it be disambiguated? If we understand *what we ought to aim at* as *what normatively explains why one ought to do something*, premiss 2 of Jackson's argument is trivially true. Furthermore, given my defence that normative explanation is essentially connected with the motivation of virtuous agents, I think that premiss 2 is substantive and may also be true when read as in (i). However, the important point, for my purposes, is that premiss 2 is definitely false if it is disambiguated as in 3. If, as Parfit claims, the formal explanation of rightness is rightness itself, then the formal explanation of rightness is not also the normative explanation. For, this would be to claim that right acts are right (normatively) because they are right, and, normative explanation, so I have suggested, does not permit reflexive explanations.

The sense of 'what one ought to aim at' then, is given by (ii) and perhaps also by (i). Having established this much, we can now ask whether premiss 1 is true on that reading. I claim that it is false, for two reasons. The first reason is the one that is admitted by Jackson himself. If we are taking 'what one ought to aim at' as a claim about what ought to motivate us, then premiss 1 is false because, as Jackson himself claims, it is rarely, if ever, the case that one ought to aim at doing right acts, except perhaps in a different,

formal sense of motivation. So, premiss 1 is false even by Jackson's own lights. But, there is a second, even more telling objection. Rightness, as intuitively understood, is a property, or relation, of acts. However, I have argued in § 2.4.1 that normative explanations can take a *wide* form, in which what normatively explains why an act is right need not be any property or relation of the act itself. For example, it may be that what normatively explains why a certain act of ϕ -ing is right is that the outcome, o , of ϕ -ing, will be best. The claim here is not that ϕ -ing is right because the outcome of ϕ -ing is best, for that explanation *does* involve a relation property of ϕ -ing. The claim is, rather, that ϕ -ing leads to a certain outcome, o , and ϕ -ing is right because o is best. Here, what normatively explains why one ought to ϕ is no property of the act of ϕ -ing at all. Consequently, there is no prospect of identifying rightness, as a property of the act of ϕ -ing, with the property that normatively explains why ϕ -ing is right. This objection is also mirrored in claims about the motivations of virtuous agents. On occasion, though surely not always, an agent should be motivated to bring about some state of affairs that is distinct from her action of bringing it about (recall Parfit's *state-given* reasons, discussed in § 2.4.1). Suppose, for example, that one ought, perhaps, to be motivated to tidy the garage. Here, it may be that the property that one ought to aim at is the tidiness of the garage, and there is certainly no prospect of identifying rightness with *that*. These considerations suffice to show that premiss 1 is false.

The same considerations can also be brought to bear against premiss 5. Premiss 5 claims that we ought to identify the property that one ought to aim at with the property that realizes rightness in the unique realization

of mature folk morality. But, as I have just argued, that claim is false. To claim that the property that realizes rightness normatively explains why right acts are right is simply to claim that right acts are right because they are right, which I deny. Furthermore, the property that realizes rightness is surely a property of acts, whereas the property that one ought to aim at need not be. So, I see no prospect that premiss 5 could be true. In summary, then, I claim that the final stage of Jackson's argument is unsound. Once its second premiss is disambiguated, each of its other premisses is shown to be false.⁵¹

4.6 Conclusion

In this chapter, I have argued that the attempts to reduce normative explanation to either metaphysical or conceptual necessity end in failure. These failures leave the defender of objective normative explanation with two significant problems.

The first problem is that the failure of the Metaphysical analysis of normative explanation is likely to rest uneasy with us unless some diagnosis is offered. My own diagnosis is that the primary flaw of the Metaphysical Analysis is that it entails that normative explanation permits the substitution of necessarily co-extensive properties within the normative explanans. Let me expand on this diagnosis. Consider again the Biconditional Law

⁵¹A further line of objection against Jackson's argument is that it would imply that there is always something that normatively explains why permissible acts are permissible. I have argued, however, that permissible acts do not always require normative explanations, because acts are, as it were, permissible by default. Compared to the other faults of Jackson's argument, however, this one is minor.

Analysis:

Biconditional Law Analysis

$E(Fx, Dx)$ iff

1. Fx and Dx ; and
2. $\Box\forall z[Fz \leftrightarrow Dz]$; and
3. it is a law that $\Box\forall z[Fz \rightarrow Dz]$.

The problem with this analysis is that in conditions 1-3, we may freely substitute property F with *any necessarily co-extensive* property, by which I mean that, the result of any such substitution does not change the truth-value of those claims. For this reason, given any property F , we may freely replace property F with any necessarily co-extensive property, such as the property of being F and an act, or the property of being F and such that $2 + 2 = 4$, or the property of being an x such that God believes that x has F . The obvious corollary of the fact that the *analysis* permits such free substitution would be that normative explanation itself permits such substitution. (The only restriction on this substitution is that property being substituted in place of F should still meet the conditions for law-likeness, but this restriction is comparatively weak.) As I have argued, however, we have a strong intuition that normative explanation does not permit this substitution. If my diagnosis of the problems of Metaphysicalism is correct, then there is a useful lesson to be drawn here, namely, that the correct analysis of normative explanation must be one that does not permit the same free substitution of necessarily co-extensive properties in the normative explanans. I draw on this lesson in the next chapter.

The second problem raised by this chapter concerns supervenience. I have argued that there is no supervenience analysis of normative explanation. However, it is almost universally accepted that there is an intimate connection between supervenience and normative explanation. How are these facts to be reconciled? If the failure of the supervenience analysis of normative explanation is taken to *deny* their intimate connection, then some error theory is surely needed to explain the universal belief in such a connection. If, on the other hand the intimate connection between supervenience and normative explanation is to be maintained, how is it to be explained, if supervenience is not part of the analysis of normative explanation? I confront this problem in the next chapter.

Five

Normative Accounts of Normative Explanation

I begin my final chapter with a brief review. My overall project in this thesis is to seek an account of *normative explanation*, the distinctively normative mode of explanation whose explananda are facts about the overall rightness or wrongness of token acts. This investigation is important given the sceptical argument that I attributed to Mackie in Chapter 2. Mackie argued that, necessarily, if there are objective moral facts, they stand in objective, normative explanatory connections to other facts. But, he insisted, there is no such thing as objective normative explanation. So, there are, after all, no objective moral facts. I have argued that we should accept the first premiss of this argument. So, if Mackie's sceptical conclusion is to be avoided, it becomes crucial to investigate objective normative explanation.

Mackie did not defend in detail the claim that there is no such thing as objective normative explanation. However, in Chapters 3–4, I rejected

accounts of normative explanation in terms of logical necessity (Inferentialism), metaphysical necessity (Metaphysicalism), and conceptual necessity (Conceptualism). Inferentialism, Metaphysicalism and Conceptualism represent three of the most credible attempts to reduce normative explanation to non-normative terms, and their failure presents a considerable problem for the moral realist, because it strongly suggests that no such non-normative reduction is possible.

In this final chapter, my argument takes a new direction. The impossibility of a *non-normative* reduction of *normative* explanation does not obviously undermine its legitimacy. If normative explanation is a genuine phenomenon, but it cannot be analysed in non-normative terms, there remains the possibility that it is *irreducibly normative*. Mackie himself would have been skeptical of this possibility, because he accepted a thoroughgoing naturalism. But, Mackie's assumptions need not be ours. Non-naturalism in ethics is a live position, and one that is worthy of consideration.

Normative explanation might be irreducibly normative in one of two, mutually incompatible, ways. Normative explanation might *itself* be irreducible (its status as normative being assumed), or normative explanation might be reducible, after all, but only to *other* normative terms that are themselves irreducible. In this chapter, I will focus on the second alternative. If, as I will argue, there is a defensible analysis of normative explanation in other normative terms, that will, in itself, constitute an argument against the claim that normative explanation is itself irreducible.

There are three main candidates to provide a reduction of normative explanation to other normative terms: normative *necessity*, normative *rea-*

sons, and normative *laws*. In § 5.1, I argue that the normative necessity analysis faces a simple but fatal objection. In § 5.2, I discuss an analysis of normative explanation in terms of normative reasons. However, I argue that this analysis is either circular or that reasons are to be further analysed in terms of normative laws. In § 5.3, I defend an analysis of *non-weighing* normative explanation in terms of normative laws, and I argue that this analysis has the virtue of explaining why overall normative properties supervene on descriptive properties. In § 5.4, I present the conclusions to my thesis and offer some suggestions for future avenues of enquiry.

5.1 Normative Necessity

My aim in this chapter is to consider accounts of normative explanation in other *normative* terms. In § 1.1.2, I suggested a general characterization of explanation as propositional connection underpinned by *necessity*. More precisely, I suggested that p explains q in some mode m only if p and q are both true, and furthermore, it is a necessity of mode m that if p then q . I claimed, then, that *necessity* is a necessary condition of explanation. This general account of explanation suggests that, if p normatively explains q , then it is *normatively necessary* that if p then q .

The idea that there is a distinctive kind of normative necessity is supported by Kit Fine. Fine writes:

It is in this [normative] sense of necessity that the moral supervenes on the natural, and indeed, such cases provide the least contentious examples of normative necessity. Suppose that D is a complete description of the world in naturalistic terms. Then

we will be inclined to make certain moral judgements about the world so described — that such-and-such a consequence was unfortunate or such-and-such an action was wrong. But in so far as we are prepared to make certain moral judgements about the world so described, we will also be prepared to say that it was no accident that they are true. In those particular circumstances, the consequences *had* to be unfortunate, the action *had* to be wrong.¹

Furthermore, Fine argues that normative necessity is itself irreducible to any other kind of necessity. Granting the availability of this irreducible normative necessity, it is natural to wonder whether it might afford an analysis of normative explanation. I will argue, however, that any such analysis must be fundamentally flawed.

As noted above, I accept that normative necessity is a *necessary* condition of normative explanation. However, to analyse normative explanation in terms of normative necessity would require that we take normative necessity to be both necessary *and* sufficient for normative explanation, and there are good reasons to deny this claim. Consider the following simple analysis:

(5.1.1) p normatively explains q iff

1. p and q ; and
2. it is normatively necessary that if p then q .

There are several familiar reasons to reject this analysis, having to do with the paradoxes of deontic logic. In particular, if the normative necessity analysis is to be effective, then the normative necessity that if p then q must

¹Fine 2002, p. 267.

somehow account for the explanatory connection between p and q . However, it would plainly be invalid to try to account for the connection between p and q in terms of the following inference: p , it is normatively necessary that if p then q ; so q . Consider, for example, a version of Chisholm's *contrary to duty* paradox.² Suppose that it is normatively necessary that, if I kill a stranger, then I ought to kill that stranger gently. But, from the fact that I *will* kill a stranger, it could hardly be inferred that I ought to kill that stranger gently. So, it is doubtful whether normative necessity can play an appropriate rôle in this analysis.

However, 5.1.1 faces a much more acute problem. One serious problem with 5.1.1 is that, within the scope of the normative necessity, and hence throughout conditions 1 *and* 2, it is surely permissible to substitute logical equivalents *salva veritate*. However, if free substitution of logical equivalents is permitted within the analysis of normative explanation, then it is also permitted within the scope of normative explanation itself. And this, I claim, must be rejected.

One reason to reject the claim that normative explanation permits substitution of logical equivalents is that it makes it impossible to avoid irrelevancies creeping in to a normative explanation. If logical equivalents can be freely substituted within normative explanations, then it follows, for example:

(5.1.2) p normatively explains q iff (p and (r or not- r)) normatively explains q .

²Chisholm 1963.

But intuitively, tautologies should not be permitted to enter explanations in this way. However, we need not rely simply on intuitions of explanatory irrelevance to reject 5.1.1, because it faces a much more serious problem. According to Davidson's³ famous slingshot argument, discussed in § 2.2.6, if a connective permits substitution of both co-referential terms and logical equivalents, then it must be a complete truth-function. I deny that normative explanation is a complete truth-function, because if it were, it would imply that either every truth normatively explains every truth, or that no truths normatively explain any truths. In short, if normative explanation were a complete truth-function, normative explanation would be either ubiquitous or impossible. This result would be catastrophic. However, if normative explanation is not a complete truth-function, then either the substitution of co-referential terms or the substitution of logical equivalents must be invalid. I have argued in § 2.2.6 that normative explanation *should* permit substitution of co-referential terms, because whether one thing objectively explains another should not depend on how we refer to the entities involved in that explanation. But, if that is so, then it follows that normative explanation does not permit substitution of logical equivalents. Consequently, every analysis of normative explanation that permits substitution of logical equivalents must be rejected, including the normative necessity analysis in 5.1.1.

I conclude, then, that despite the favourable picture afforded by Fine's claim that there is an irreducible variety of normative explanation, it can nevertheless provide no analysis of normative explanation. I turn now to

³Davidson 1967; Davidson 1969.

the alternative possibility of analysing normative explanation in terms of normative reasons.

5.2 Normative Reasons

In this section, I consider whether normative explanation might be analysed in terms of normative reasons. I begin by clarifying the concept of a *normative reason*. I assume that normative reasons are (or are provided) by *facts*. Very roughly, a fact is (or provides) a normative reason just in case it *counts in favour* or *counts against* something, typically an act or mental state.⁴ Roger Crisp has argued for a further distinction between *grounding* and *justificatory* reasons.⁵ Roughly put, *grounding reasons* count in favour of actions while *justificatory reasons* justify agents in performing actions. In ideal cases, these reasons may coincide, but in less ideal cases, when, for example, the agent has imperfect knowledge of what grounding reasons there are, whether or not his action is justified is determined by a different set of considerations, the *justificatory* reasons. Henceforth, by *normative reasons* I will mean *grounding reasons*.

Normative reasons are most often understood to be *pro tanto* or, in other terminology, *contributory*. Dancy characterizes this idea, writing:

A contributory reason for action is a feature whose presence makes something of a case for acting, but in such a way that the overall case for doing that action can be improved or strength-

⁴Normative reasons thus differ from motivating reasons, which are to be defined in terms of their rôle within motivation and the explanation of action. I discuss this distinction further in Appendix A.

⁵Crisp 2006, pp. 39–40.

ened by the addition of a second feature playing a similar rôle. Also, a contributory reason on one side is not necessarily destroyed by the presence of a reason on the other side.⁶

I follow Broome and others⁷ in taking this conception of reasons to be best understood as a *force* model. On this view, normative reasons are understood to be (or to give rise to) complexly interacting normative forces. The force model of *pro tanto* normative reasons helps to explain a number of their distinctive features. On the force model, normative forces are associated with particular direction and magnitude. A fact may count *strongly* in favour of an act or *weakly* against. Furthermore, there can be, simultaneously, a plurality of normative forces counting in favour of an act, and a plurality counting against. Finally, the overall normative status of any action is somehow determined by ‘weighing up’ these various forces, but in a way that is complicated by the fact that normative reasons interact in complex ways.

In this section, I consider whether *pro tanto* reasons can offer an analysis of normative explanation. Jonathan Dancy has recently suggested that there may be a class of *enticing* reasons that, although they are genuinely normative, are too weak to explain an obligation. Dancy writes:

Enticing reasons are to do with what would be fun, amusing, attractive, exciting, pleasant, and so on. They can be stronger and weaker, and they are often strong enough for action. But [...] they never take us to an ought; it is not true of an enticing reason that if one has one of them and no reason of any other sort, one ought do to what the reason entices one to do.⁸

⁶Dancy 2004*b*, p. 15.

⁷Pietroski 1993; Broome 2004*a*; Broome 2008, Ch. 3.

⁸Dancy 2004*b*, p. 21. See also Dancy 2004*a*.

Since I am only concerned with reasons that *can* explain an obligation, I shall henceforth ignore enticing reasons.

Let us consider how normative explanation might be analysed in terms of *pro tanto* normative reasons. I take the driving intuition behind a reasons analysis of normative explanation to be the thought that an act x is obligatory iff there is overall reason to do x . We might take this to suggest the following analysis of normative explanation:

(5.2.1) The fact that x has properties F_1, \dots, F_i normatively explains that x is obligatory iff the fact that x has F_1, \dots, F_i provides overall reason to x .

5.2.1 faces two problems. The first problem is that it uses the concept of *overall reason* rather than the concept of a *pro tanto* reason. The second problem is that whether there is *overall* reason to do an act is not determined solely by the considerations that count in favour of it. Whether there is *overall* reason to do an act is determined by whether the reasons in favour of doing it outweigh the reasons against. It seems more natural, however, to take the obligation to do an act to be explained only by the *pro tanto* reasons in favour of doing it. So, I suggest this amendment:

(5.2.2) The fact that x has properties F_1, \dots, F_i normatively explains that x is obligatory iff

1. F_1, \dots, F_i are all the *pro tanto* reasons in favour of x ; and
2. the *pro tanto* reasons in favour of x together outweigh the *pro tanto* reasons against x .

I take 5.2.2 to be a plausible proposition, provided that it is understood as nothing more than a necessary and sufficient condition of normative explanation. However, if 5.2.2 is genuinely to provide an *analysis*, then it must meet two further conditions. First, if 5.2.2 is not to be circular, then normative reasons must not themselves need to be understood in terms of normative explanation. Second, if 5.2.2 is to provide an analysis of normative explanation in *irreducible* normative terms, then *pro tanto* reasons must not themselves be analysable in terms of other normative phenomena, such as normative laws. Doubts arise over both of these conditions.

If 5.2.2 is to provide an analysis of normative explanation in terms of *pro tanto* normative reasons, then that analysis must not itself need to be understood in terms of normative explanation. One reason to doubt whether this condition can be met comes from John Broome. Broome argues that *pro tanto* normative reasons are themselves to be understood via the concept of normative explanation:

I have described the characteristic role that *pro tanto* reasons play in an explanation of why you ought to ϕ : each has a weight, and the fact that you ought to ϕ is explained by the fact that the reasons for you to ϕ outweigh the reasons for you not to ϕ according to an aggregating function. A *pro tanto* reason is defined as a fact that plays this characteristic role in a weighing explanation.⁹

Broome argues that *pro tanto* reasons can be identified by their characteristic rôle in a particular kind of explanation of normative facts, which he calls a *weighing* explanation. I assume that *weighing* explanation of an

⁹Broome 2004a, p. 38.

normative fact is in the normative mode. If Broome's analysis of *pro tanto* reasons is correct, then it would, of course, make 5.2.2 circular. *Pro tanto* reasons cannot provide an analysis of normative explanation if *pro tanto* reasons are themselves to be analysed in terms of normative explanation.

There is a second reason to think that 5.2.2 might be circular, even if *pro tanto* reasons cannot be defined by their rôle in a weighing explanation. In addition to the concept of *pro tanto* normative reasons, 5.2.2 includes the concept of one set of *pro tanto* reasons *outweighing* another set. If the 5.2.2 is not to be circular, then some account is needed of this concept that does not make essential reference to a *weighing* normative explanation. It is questionable, however, whether this can be achieved. For example, one obvious way to define what it is for a first set of reasons to outweigh a second set, is for the first set of reasons to provide a successful weighing explanation of a normative fact. This account of *outweighing*, however, would obviously make 5.2.2 circular. One might try to avoid this difficulty by amending condition two of 5.2.2, so as to read:

- 2'. the *pro tanto* reasons in favour of *x* are together *stronger* than the *pro tanto* reasons against *x*.

In order for this amendment to remove the threat of circularity, some account is needed of what it is for one set of reasons to be *stronger* than another that does not make essential reference to a weighing normative explanation. Again, I think it is questionable whether this could be achieved.

I have argued that it is questionable whether 5.2.2 provides a non-circular analysis of normative reasons. I will now put those doubts asides

in order to focus on a different issue. Let us assume, for the sake of argument, that 5.2.2 provides a non-circular analysis of normative explanation. The question now arises whether 5.2.2 analyses normative explanation down to fundamental terms. In particular, does the analysis stop at normative reasons, or can normative explanation be analysed still further? In recent literature, it has been argued that normative reasons are irreducible. For example, Scanlon writes:

I will take the idea of a reason as primitive. Any attempt to explain what it is to be a reason seems to me to lead back to the same idea: a consideration that counts in favor of it. “Counts in favor how?” one might ask. “By providing a reason for it” seems to be the only answer.¹⁰

Scanlon’s argument relies on the claim that the only answer to the question of how a consideration counts in favour of a course of action is by providing a reason for it. We have already seen one objection to this claim, namely, Broome’s claim that *pro tanto* normative reasons are to be defined via their rôle in *weighing* normative explanations. However, even putting this possibility aside, I find Scanlon’s claim puzzling, because there is a readily available alternative conception. According to a long established, if perhaps implicit, view, what it is for a consideration to count in favour of an act is simply for that consideration and that action to be subsumed under the antecedent and consequent of a normative *law*. For example, the fact that one promised to ϕ provides a reason to ϕ iff that fact is subsumed under the normative law that promises ought to be kept. Call this the *Nomological Conception* of normative reasons.

¹⁰Scanlon 1998, p. 17. The irreducibility of reasons is also defended by Parfit (2008, Ch. 1)

The Nomological Conception of *pro tanto* normative reasons faces a number of acknowledged difficulties, but they may not be insuperable. A first difficulty, for example, is that *pro tanto* reasons can be defeated. However, recent work in ethics and in the philosophy of science has emphasized a distinction between *strict* and *defeasible* laws, and some have claimed that defeasible laws are predominant in ethics.¹¹ If there are defeasible laws, then the Nomological Conception would seem capable of accommodating defeasible normative reasons. A second source of difficulty for the Nomological Conception might seem to derive from the thought that *pro tanto* reasons are often understood by analogy with *forces*. In order to cope with this difficulty, we need a further distinction between *force* and *non-force* laws. However, the idea of force laws is already well established in the natural sciences, and there is no obvious reason why moral laws could not be conceived as *force* laws in an analogous way.¹² I return to this discussion of normative laws in § 5.3.

I defend the idea that normative reasons must be understood in terms of laws because, without laws, normative reasons cannot account for the near-universally accepted claim that the normative supervenes on the descriptive. Contrary to my position, Jonathan Dancy has argued that considerations of normative reasons can indeed account for this supervenience claim, so it will be instructive to see where this argument breaks down. Dancy writes:

¹¹Armstrong 1983, Ch. 10; Cartwright 1983, Ch. 2; Pietroski 1993; Pietroski and Rey 1995; Little 2000; Lance 2004; Lance and Little 2006*b*; Lance and Little 2006*a*; Bird 2007, § 3.3; Lance and Little 2007.

¹²The conception of ethics as dealing in force laws akin to those of physics is skillfully developed by Pietroski 1993; Pietroski and Rey 1995. For a recent attempt to spell out the idea of a force law in physics, see Rupert 2008.

What then explains the fact that moral properties supervene on the non-moral? My own view is that this is a consequence of the fact that they result from the non-moral. The properties from which the wrongness of this action results are the reasons why it is wrong, the ground for its wrongness. Now we know already that if we move to another case holding just these properties fixed, we may yet get an action that is not wrong; differences elsewhere may conspire to prevent the original wrong-making properties from doing that job in the new case. But if we move to another case holding fixed all the non-moral properties of the case whatever, we know in advance that no conspiracy of that sort can happen. Whatever was a reason in the first case must remain a reason in the second, and nothing that was not a reason in the first can become a reason in the second. So necessarily, the reasons must remain the same, and if so, their rational result must remain the same. So if the first act is wrong, the second one must be wrong too.¹³

I take this argument to fail because it conflates two quite different claims. Let us suppose that there are two descriptively indistinguishable acts, a and b , in two possible worlds, w_1 and w_2 , and let us suppose that a is wrong because it has a descriptive property, F . Since a and b are descriptively indistinguishable, b also has property F . Furthermore, a and b must be exactly alike in terms of the presence or absence of those properties that, in a 's case, function either as *reasons*, *enabling* conditions or *disabling* conditions. Dancy's first claim is that, if b had property F , but differed from a in respect of those properties that function, in a 's case, as enabling or disabling conditions, then property F might not provide a reason against b , and hence might not explain why b is wrong. This claim certainly seems plausible, and I will accept it for the sake of argument. However, from this first claim Dancy infers the quite different claim that if b has precisely

¹³Dancy 2004b, p. 89.

the same properties as *a*, including the property, *F*, that explains why *a* is wrong, alongside the properties that function, in *a*'s case, as enabling or disabling conditions, then *b*'s having *F* must also provide a reason against *b*, and so must also explain why *b* is wrong. But, this second claim does not at all follow from the first. That *different* reasons might arise for *a* and *b* if they *differ* in properties, does not entail that the *same* reasons must arise for *a* and *b* if they are have precisely the *same* properties. To assume, as Dancy does, that a set of properties must function in the same way whenever it is instantiated, is simply to assume what needs to be shown, and hence to beg the question. Now, there is an obvious way to defend the claim that a set of properties must function in the same way whenever it is instantiated, and that is to claim that those properties function that way as a matter of *law*. But, if normative reasons are irreducible, no recourse to this explanation is possible. There is, then, no argument from irreducible normative reasons to the supervenience of the normative on the descriptive.

I take the failure of the argument from normative reasons to supervenience to be a serious objection to the reasons analysis, because it leaves the supervenience claim looking objectionably mysterious. By contrast, I do take there to be an argument from normative *laws* to supervenience, and I develop this argument in the next section.

5.3 A Nomological Analysis of Normative Explanation

In this section, I develop a partial analysis of normative explanation in terms of metaphysically irreducible normative laws. More precisely, I will analyse *non-weighing* normative explanations. I call this the *Nomological Analysis*. I will say a little more about the contrast between *weighing* and *non-weighing* explanations below. I will not attempt an analysis of *weighing* explanations because they present considerable extra difficulties which I do not know how to accommodate.

The importance of normative laws to an analysis of normative explanation was already suggested by the Inferentialist and Metaphysicist accounts I considered in Chapters 3–4. Although normative laws were not strictly speaking essential to those accounts, the addition of normative laws to those analyses greatly increased their plausibility. In this section, I build on the idea that normative laws are an essential component in the analysis of normative explanation, but I add the further idea that such laws are metaphysically primitive.

To claim that a certain entity is a metaphysical primitive is not in itself philosophically disreputable. For example, Williamson¹⁴ has recently argued that knowledge is a metaphysical primitive. There are, however, difficulties in *defending* the claim that a certain entity is metaphysically primitive. I take this, however, to be primarily a methodological difficulty. If an entity cannot be analysed into other, less contentious components,

¹⁴Williamson 2000.

what arguments can be offered in defence of its legitimacy? In Chapter 2, I offered two ways of defending the legitimacy of normative explanation. The first was by characterizing its *formal* properties, and thereby showing at least that the phenomenon is not confused and that we have a reasonable concept of it. This offers some, albeit weak, support for the claim that normative explanation is a genuine phenomenon. A second way of defending the legitimacy of a metaphysical primitive is through its connection with *non-primitive* phenomena. Roughly speaking, this is the method of reductive analysis in reverse. Instead of defending a phenomenon by reducing it to more primitive metaphysical constituents whose legitimacy is not in doubt, one aims to show that one or more complex phenomena, whose legitimacy is not in doubt, can be explained in terms of *it*. I call this the *constructive* method of analysis, and its explanatory power is increased significantly if the more complex phenomena cannot otherwise be explained. In this section, I offer a *constructive* defence of normative laws by arguing that they offer the best, perhaps the *only*, realist explanation of why normative properties supervene on descriptive properties.

5.3.1 Normative Laws

I proceed now to develop the idea of irreducible normative laws. I will assume a primitive propositional operator, ‘It is a normative law that *p*’. Michael Tooley¹⁵ has suggested that *causal* laws might be represented as ‘It is a law that *As* cause *Bs*’, with the concept of a *cause* appearing within the scope of the law. A similar move on my part would lead to laws of the

¹⁵Tooley 2003.

form ‘It is a normative laws that F s normatively explain D s.’ However, this conception of normative laws would, of course, provide no reductive analysis of normative explanation. Consequently, I shall adopt the more standard view that laws govern generalizations in which explanatory concepts do not appear.

I will take the operator, ‘It is a normative law that p ’ to govern universal generalizations that meet the conditions for law-likeness previously discussed in § 3.2. In addition, I will assume that normative laws govern generalizations between a *descriptive* property and an property of overall *normative* status, such as being overall obligatory, permissible, or wrong. For simplicity of presentation, I will assume that laws are generalizations over the properties of a single entity. I assume that is, that ‘ p ’ is to be a generalization of the form $\forall x[Fx \rightarrow Dx]$, where F is a descriptive property and D is an overall normative property. However, the Nomological Analysis that I will develop can, in principle, be extended to accommodate for generalizations of the form $\forall x[Fx \rightarrow \exists y Dy]$.

In developing my Nomological Analysis, I will make use of two important distinctions concerning laws. The first is the distinction between *strict* laws and *defeasible* laws.¹⁶ A strict law can be characterized as one that implies its generalization. That is to say, it is a *strict* law that $\forall x[Fx \rightarrow Dx]$ only if that generalization is true, i.e. only if $\forall x[Fx \rightarrow Dx]$. By contrast, defeasible laws need not imply their generalizations. Put another way, strict laws are

¹⁶Armstrong 1983, Ch. 10; Cartwright 1983, Ch. 2; Pietroski 1993; Pietroski and Rey 1995; Little 2000; Lance 2004; Lance and Little 2006*b*; Lance and Little 2006*a*; Bird 2007, § 3.3; Lance and Little 2007.

factive and defeasible laws are not.¹⁷ The non-factiveness of defeasible laws has led to considerable controversy regarding their nature, with some arguing that defeasible laws simply collapse into triviality. For the sake of argument, however, I will assume that defeasible laws are not trivial. In discussing defeasible laws, however, I will make use of the idea that a given law may *apply*, or *fail to apply*, to a particular case. More precisely, I will assume that, if the law that $\forall x[Fx \rightarrow Dx]$ *applies* to a particular entity, y , then $[Fy \rightarrow Dy]$.

The second distinction in laws is between what we might call *force* laws and *non-force* laws. Sometimes this distinction is assimilated to the distinction between defeasible and strict laws, but I would argue that there are two different phenomena here. On a simple view, a non-force law is a law that directly governs the overall behaviour of a system. A force law, by contrast, helps to govern the behaviour of a system, perhaps in combination with other force laws, by governing the manifestation of forces within that system. The overall behaviour of the system is then somehow determined by all the various forces that are manifested within it. The distinction between the way in which force and non-force laws explain the overall behaviour of a system is brought out by John Broome's metaphor of *weighing* versus *non-weighing* explanations. Broome characterizes *weighing* explanations by reference to a loose analogy with *mechanical weighing*:

[...] when the fact that you ought to ϕ is explained by *pro tanto* reasons, the explanation retains central elements of the mechanical analogy. It includes one or more reasons for you to ϕ , and

¹⁷The claim that a defeasible law is non-factive does not imply that the generalization of a defeasible law is false. In some cases, a defeasible law may happen to go undefeated.

it may also include reasons for you not to ϕ . These reasons are analogous to the objects in the left-hand and right-hand pans of the scales. Each reason is associated with a metaphorical weight. This weight need not be anything so precise as a number; it may be an entity of some vaguer sort. The reasons for you to ϕ and those for you not to ϕ are aggregated or weighed together in some way. The aggregate is some function of the weights of the individual reasons. The function may not be simply additive, as it is the mechanical case. It may be a complicated function, and the specific nature of the reasons may influence it. Finally, the aggregate comes out in favour of your ϕ ing, and that is why you ought to ϕ .¹⁸

Understanding the behaviour of force laws, and the *weighing* explanations that go with them, present considerable extra problems that considerations of space prevent me from developing. For this reason, in developing my Nomological Analysis, I will permit normative laws to be *strict* or *defeasible*, but I will stipulate that they are to be non-force laws. One consequence of this restriction is that my Nomological Analysis is not a complete account of normative explanation. It is instead a restricted model that relies on an idealising assumption that there are no force laws, and hence no *pro tanto* normative reasons. Nevertheless, I think my Nomological Analysis is a promising model of *non-weighing* normative explanation.

One notable feature of the conceptions of normative laws employed by the Inferentialist and Metaphysicist accounts was that it took normative laws to be metaphysically necessary. The idea that normative laws (by which I mean, *fundamental* normative laws) are metaphysically necessary has much intuitive plausibility, and I, too, will make this assumption. However, I suggest that normative laws are necessary in a way that differs from

¹⁸Broome 2004a, p. 37.

the previous accounts that we have seen. Previously, I have taken the idea that normative laws are necessary to mean that they govern necessary generalizations i.e. ‘It is a normative law that $\Box\forall x[Fx \rightarrow Dx]$ ’. However, if laws are a metaphysically entity, I suggest that their necessity is better conceived so as to attach to the *law*, and not to the generalization. I will assume, then, that normative laws are necessary in the sense that, if it is possible that it is a normative law that $\forall x[Fx \rightarrow Dx]$, then it is necessary that it is a normative law that $\forall x[Fx \rightarrow Dx]$.

5.3.2 A Nomological Analysis of Non-Weighing

Normative Explanation

With these clarifications and assumptions in place, my *Nomological Analysis* of *non-weighing* normative explanation can now be presented:

(5.3.1) Fx normatively explains Dx iff:

1. Fx and Dx ; and
2. it is a (strict or defeasible) normative law that $\forall y[Fy \rightarrow Dy]$; and
3. the (strict or defeasible) normative law that $\forall y[Fy \rightarrow Dy]$ applies to x .

Condition 1 of this analysis expresses the requirement of factiveness. Strictly speaking Dx may be redundant here, but its inclusion does no harm. Condition 2 is the nomological component of the analysis. Condition 3 allows

for the possibility of defeasible laws, whose application to any given case cannot be taken for granted.

If the Nomological Analysis of normative explanation is to be defensible, it must avoid the mistakes that I have diagnosed in other analyses. One virtue of 5.3.1 is that it does not take laws to be part of the normative *explanans*. In this way, 5.3.1 avoids the mistake made by the Inferentialist analysis. To avoid the other mistakes that I have diagnosed, however, further constraints must be imposed on the conception of irreducible normative laws. One obvious constraint on the success of 5.3.1 is that, if the dire consequences of the slingshot argument are to be avoided, then it must not be permitted to substitute logical equivalents within the scope of the normative law operator. I defended that constraint in § 5.1. A second constraint on 5.3.1, is that it must not permit the substitution of metaphysically necessary equivalents in the explanans. We saw in Chapter 4 that the Metaphysical analysis of normative explanation could entail that:

(5.3.2) Fx normatively explains Dx iff (Fx and x is an act) normatively explains that Dx .

To avoid this difficulty, we must deny that the property in the antecedent of a normative law's generalisation may be freely substituted with any necessarily co-extensive property.

I have now presented an analysis of non-weighting normative explanation in terms of irreducible normative laws, and I have argued that those laws must meet certain conditions if they are to avoid the mistakes of earlier analyses. Of course, it is one thing to develop a conception of normative

laws that will fulfil a certain argumentative rôle, it is another to suppose that there really are normative laws of that kind. I defend the claim that there are irreducible normative laws by arguing that positing normative laws of this kind offers the best realist explanation of the near-universally accepted claim that normative *strongly supervenes* on the descriptive. I proceed now to develop this argument.

5.3.3 From Normative Explanation To Deontic Supervenience

In this subsection, I will defend the claim that there are irreducible normative laws by arguing that this supposition offers a realist explanation of why the normative supervenes on the descriptive. More precisely, I will defend a restricted version of this claim, namely that the *overall* normative properties of obligation, permissibility and wrongness supervene on descriptive properties. It is often argued that moral realism, especially non-naturalist moral realism, cannot explain the supervenience of the normative on the description, and so leaves it looking objectionably mysterious.¹⁹ I argue, to the contrary, that normative laws permit just such a realist explanation.

My argument from normative laws to supervenience is important for a second reason. In Chapter 4, I argued against the supervenience *analysis* of normative explanation. I argued, more precisely, that the analysis was insufficient and perhaps also unnecessary. However, many philosophers have supposed that there is a necessary connection between supervenience and

¹⁹Blackburn 1984, pp. 182–190; Blackburn 1985; Horgan and Timmons 1992; Sobel 2001.

normative explanation. Michael Smith writes:

The claim that the evaluative supervenes on the natural divides into two parts. The first part is the claim that the evaluative is *supervenient*; the second part is a claim about what the evaluative supervenes on — namely, the *natural*. As regards the first part, the relevant fact seems to be that it is simply incoherent to suppose that evaluative claims could be *barely true*. Evaluative claims must always be *made true* by other claims. [...] This is all that is meant to say that the evaluative is supervenient.²⁰

This leaves a problem. If normative explanation cannot be analysed in terms of supervenience, why suppose they are necessarily connected? The claim that supervenience has no connection with normative explanation would be counterintuitive enough to require an error theory. If, on the other hand, the connection is allowed to stand, some explanation of it is surely needed. Yet, what explanation can be offered, once the supervenience analysis of normative explanation is denied? I propose to defend the idea that there is a necessary connection between normative explanation and supervenience, but that the direction of explanation runs in the opposite direction to what is usually supposed. Rather than analysing normative explanation in terms of supervenience, I argue *from* normative explanation *to* supervenience.

In arguing for the supervenience of normative properties on descriptive properties, let me note one favourable feature of the situation. Supervenience is a transitive relation. If set \mathbb{A} supervenes on set \mathbb{B} , and \mathbb{B} on \mathbb{C} , then \mathbb{A} supervenes on \mathbb{C} . It is generally accepted that, necessarily, an act is permissible iff it is neither obligatory nor wrong. It follows that *permissibility* already supervenes on obligation and wrongness. So, in arguing for

²⁰Smith 2000, p. 108.

the supervenience of overall normative properties on descriptive properties, I need only argue that obligation and wrongness supervene. If obligation and wrongness supervene on descriptive properties, so, too, does permissibility. Consequently, my aim is to show that considerations of normative explanation entail the following claim:

Deontic Supervenience Necessarily, for all x , if x is obligatory (or wrong), then there is some descriptive property, F , such that x has F and necessarily if any y has F , then y is obligatory (or wrong).

In what follows, I will refer to obligation and wrongness as the *special* deontic properties.

The idea that considerations of normative explanation entail that normative properties supervene on descriptive properties is suggested by both Sidgwick and Hare. Sidgwick writes:

In the variety of coexistent physical facts we find an accidental or arbitrary element in which we have to acquiesce [...]. But within the range of our cognitions of right and wrong, it will generally be agreed that we cannot admit a similar unexplained variation. We cannot judge an action to be right for A and wrong for B , unless we can find in the natures or circumstances of the two some difference which we can regard as a reasonable ground for difference in their duties.²¹

Hare's argument concerns the aesthetic properties of two identical pictures, P and Q, which are taken to supervene on its nonaesthetic properties, but it is easily extended to the moral case. Hare writes:

²¹Sidgwick 1962, pp. 208-9.

[...] there is one thing we cannot say; we cannot say ‘P is exactly like Q in all respects save this one, that P is a good picture and Q not’. If we were to say this, we should invite the comment, ‘But how can one be good and the other not, if they are exactly alike? There must be some *further* difference between them to make one good and other not.’ Unless we at least admit the relevance of the question ‘What makes one good and the other not?’ we are bound to puzzle our hearers [...].²²

Sidgwick claims that deontic properties supervene on descriptive properties because it would be irrational to admit ‘unexplained variations’ in overall deontic status for which there is no ‘reasonable ground’. Similarly, Hare claims that to deny the supervenience of aesthetic properties would be to leave oneself with no answer to the question of what makes it the case, that is, what *explains*, that one painting is good and the other not. The striking element of both these passages is that the supervenience of moral or aesthetic properties on descriptive properties is defended on the grounds that to deny it would imply an explanatory absurdity. Thus, supervenience is taken to follow from considerations of explanation. However, as suggestive as these passages are, they do not yet amount to a perspicuous argument. I aim to fill the explanatory gap.

I will present two versions of my argument from normative explanation to Deontic Supervenience. The first argument proceeds under a simplifying assumption, but demonstrates my general line of thought. The second argument proceeds without the simplifying assumption. A key premiss in both versions of my argument is that normative laws are *metaphysically necessary*. That is to say, I assume that a normative law holds in *all* possible

²²Hare 1986, p. 81.

worlds if it holds in *any*. A second key premiss in my argument is a version of the claim of *explanatory moral dependence*, which I discussed in § 2.4.1. In particular, I will assume that whenever an act is obligatory or wrong, then that is to be explained by its having some descriptive property. In § 2.4.1, I argued that this assumption was questionable. However, this version of explanatory moral dependence is implied by my Nomological Analysis, in combination with my assumption that normative laws connect descriptive with normative properties.

I begin with a simplified version of my argument in which I assume that all normative laws are *strict*.

1. (Premiss) For all acts x and special deontic properties D , if Dx then there exists descriptive property, F , such that Fx normatively explains Dx .
2. (Premiss) For all acts x , descriptive properties F and special deontic properties D , if Fx normatively explains Dx , then it is a strict normative law that $\forall y[Fy \rightarrow Dy]$.
3. (From 1, 2) For all acts x and special deontic properties D , if Dx then there exists descriptive property, F , such that it is a strict normative law that $\forall y[Fy \rightarrow Dy]$.
4. (Premiss) For all descriptive properties F and special deontic properties D , if it is a strict normative law that $\forall y[Fy \rightarrow Dy]$ then it is metaphysically necessary that it is a strict normative law that $\forall y[Fy \rightarrow Dy]$.

5. (From 3, 4) For all acts x and special deontic properties D , if Dx then there exists a descriptive property, F , such that it is metaphysically necessary that it is a strict normative law that $\forall y[Fy \rightarrow Dy]$.
6. (Premiss) For all descriptive properties F and deontic properties D , if it is a strict normative law that $\forall y[Fy \rightarrow Dy]$ then $\forall y[Fy \rightarrow Dy]$.

Therefore:

7. For all acts x and special deontic properties D , if Dx then there exists a descriptive property, F , such that it is metaphysically necessary that $\forall y[Fy \rightarrow Dy]$.

This argument may be understood as follows. Premiss 1 is the claim of Explanatory Moral Dependence discussed above. Premiss 2 comes from my Nomological Analysis of normative explanation. Premiss 4 is the claim, introduced above, that normative laws are metaphysically necessary. Premiss 6 is the claim, also argued above, that strict laws imply their generalizations. I take each premiss, and hence also the conclusion, to hold with metaphysical necessity. 7 is thus the claim of Deontic Supervenience, which was to be proven. This argument shows one way in which Deontic Supervenience can follow from considerations of normative explanation.

My aim now is to relax my simplifying assumption that all normative laws take a strict form, and hence to show how my argument can be extended to cover the case of *defeasible* laws. The key idea in this extension of my argument is that, although the application of a defeasible law to a particular case cannot be taken for granted, nor is it a matter of chance.

More precisely, my key premiss is that whether a defeasible normative law applies to a given act *strongly supervenes* on the descriptive properties of that act. To see how this assumption works in practice, I begin with an informal presentation of my argument. Let us suppose that we have two descriptively indistinguishable acts, *a* and *b*, which may be in different possible worlds. Assume that act *a* is obligatory. If *a* is obligatory, then *a* has some descriptive property, *F*, such that *Fa* explains why *a* is obligatory. By my Nomological Analysis of normative explanation, if *Fa* explains why act *a* is obligatory, then there is a normative law to the effect that all acts with property *F* are obligatory, and furthermore, that law *applies* to *a*. Given the assumption that normative laws are metaphysically necessary, it follows that it is metaphysically necessary that it is a law that all acts with property *F* are obligatory. Now let us turn our attention to *b*. *Ex hypothesi*, *a* and *b* are descriptively indistinguishable. So, given that *a* has property *F*, so, too, does *b*. Next, given that the law that all acts with property *F* are obligatory holds in all possible worlds, it also holds in *b*'s world. This is not yet quite enough, however, to ensure that that law *applies* to *b*. To establish this last step, note that *a* and *b* are descriptively indistinguishable, and the law that all acts with property *F* are obligatory *does* apply to *a*. That being so, the law that all acts with property *F* are obligatory must also apply to *b*, because *b*, being descriptively indistinguishable from *a*, shares all the same enabling and disabling conditions as *a*. Finally, if it is true *both* that *b* has property *F* *and* that the law that all acts with property *F* are obligatory *applies* to *b*, then *b* must also be obligatory, and this yields the desired conclusion. The key premiss in my argument is the claim that the

applicability of a normative law to any particular case *strongly supervenes* on the descriptive properties of that case. Hence, if we have one case where a defeasible law applies, then it must also apply in all descriptively identical cases. This is the assumption that I take to be implicit in the arguments of Hare and Sidgwick: given one case where a law applies, it would be absurd to suppose that the law did not apply in a descriptively indistinguishable case. I now proceed to a more formal presentation of my argument.

1. (Premiss) For all acts x and special deontic properties D , if Dx then there exists a descriptive property, F , such that Fx normatively explains Dx .
2. (Premiss) For all acts x , descriptive properties F and special deontic properties D , if Fx normatively explains Dx then it is a normative law that $\forall z[Fz \rightarrow Dz]$, and the law that $\forall z[Fz \rightarrow Dz]$ applies to x .
3. (From 1, 2) For all acts x and special deontic properties D , if Dx then there exists a descriptive property, F , such that Fx , and it is a normative law that $\forall z[Fz \rightarrow Dz]$, and the law that $\forall z[Fz \rightarrow Dz]$ applies to x .
4. (Premiss) For all properties F and D , if it is a normative law that $\forall z[Fz \rightarrow Dz]$ then it is metaphysically necessary that it is a normative law that $\forall z[Fz \rightarrow Dz]$.
5. (From 3, 4) For all acts x and special deontic properties D , if Dx then there exists a descriptive property, F , such that Fx and it is

metaphysically necessary that is a normative law that $\forall z[Fz \rightarrow Dz]$, and the law that $\forall z[Fz \rightarrow Dz]$ applies to x .

6. (Premiss) For all acts x , descriptive properties F and special deontic properties D , if it is law that $\forall z[Fz \rightarrow Dz]$ and that law applies to x , then there exists a descriptive property, F' , such that $F'x$ and it is metaphysically necessary that for all acts y , if $F'y$ and it is a law that $\forall z[Fz \rightarrow Dz]$, then the law that $\forall z[Fz \rightarrow Dz]$ applies to y .
7. (From 5, 6) For all acts x and special deontic properties D , if Dx , then there exist two descriptive properties, F and F' , such that:
 - (i) Fx , and it is metaphysically necessarily that it is a law that $\forall z[Fz \rightarrow Dz]$; and
 - (ii) $F'x$, and it is metaphysically necessary that for all acts y , if $F'y$ and it is a law that $\forall z[Fz \rightarrow Dz]$, then that law applies to y .
8. (Premiss) For all acts x , descriptive properties F and special deontic properties D , if it is law that $\forall z[Fz \rightarrow Dz]$ and that law applies to x , then $[Fx \rightarrow Dx]$.
9. (From 7, 8) For all acts x and special deontic properties D , if Dx , then there exist two descriptive properties, F and F' , such that Fx and $F'x$, and it is metaphysically necessary that $\forall z[F'z \rightarrow [Fz \rightarrow Dz]]$ (and, hence, by trivial rearrangement, it is metaphysically necessary that $\forall z[[F'z \wedge Fz] \rightarrow Dz]$).

Therefore:

10. (From 9) For all acts x and special deontic properties D , if Dx then there is a (perhaps conjunctive) descriptive property, F , such that Fx and it is metaphysically necessary that $\forall z[Fz \rightarrow Dz]$

This argument can be understood as follows. Premiss 1 is the claim of Explanatory Moral Dependence discussed above. Premiss 2 expresses the claim that normative explanation is backed by normative laws, as implied by my Nomological Analysis. Premiss 4 is the assumption that normative laws are metaphysically necessary. Premiss 6 is the key claim, defended above, that the application of a law to a particular act strongly supervenes on the descriptive properties of that act. I express this supervenience claim using the standard *modal operator* version of strong supervenience discussed in § 4.4. Premiss 8 expresses the intuitive consequence of a law's *applying* to a given case. 9 might easily itself be taken to express a claim of supervenience, but on the assumption that we can form conjunctive properties, 9 entails the more usual formulation of supervenience in 10. I take each premiss, and hence also the conclusion, to hold with metaphysical necessity. Hence, 10 is equivalent to Deontic Supervenience, which was to be proved.

This argument shows how considerations of normative explanation can entail Deontic Supervenience even if some, or all, normative laws take a defeasible form. Some might object that my argument is trivial because the claim of Deontic Supervenience is covertly contained among its premisses. But, I reject this charge. My arguments premisses *contain* its conclusion only in the way that all formally valid arguments do. I take my argument to be sound and non-trivial, and its success counts strongly in favour of the

Nomological Analysis of normative explanation, and its assumption that there are irreducible normative laws.

In this section, I have defended an analysis of *non-weighing* normative explanation in terms of irreducible normative laws, which may be strict or defeasible. One obvious direction for future research is to try to extend this analysis to incorporate *weighing* explanations. In order to achieve that, my analysis would need to be extended to accommodate what I have called normative *force* laws, and some account would need to be offered of how the various *pro tanto* reasons — understood as those considerations that fall under the antecedent of a normative force law that applies to the current situation — combine to determine the overall normative status of an act. If the Nomological Analysis can be extended to accommodate *pro tanto* reasons, then it may also be possible to extend my argument from normative laws to normative supervenience to accommodate them. In this way, one might hope to use normative laws to explain not, just the supervenience of *overall* normative properties, but also the supervenience of *pro tanto* reasons. I leave this task to future research.

5.4 Conclusion To Thesis

In this thesis, I have sought an objective account of normative explanation. My investigation was motivated, in particular, by a desire to find a realist response to the sceptical challenge posed by Mackie's *argument from explanatory queerness*. In Chapters 3–5, I developed accounts of normative explanation in terms of logical, metaphysical and conceptual necessity, but

I argued that each of them was vulnerable to fatal objection. The failure of these accounts demonstrates the extraordinary power of Mackie's sceptical argument, and the difficulty it poses for the defender of moral realism.

In this final chapter, I have developed a realist response to Mackie's argument in terms of irreducible normative laws. I have defended an analysis of *non-weighing* normative explanation in terms of those laws, and I have argued that there is good reason to posit those laws because they offer what is, perhaps, the only realist explanation of why the normative supervenes on the descriptive. I concede that my non-naturalist proposal is unlikely to persuade a committed naturalist, but I take non-naturalism in ethics to be a live option, and one whose benefits are demonstrated by my supervenience argument.

In this thesis, I have been able only to scratch the surface of the issues raised by a realist response to Mackie's sceptical challenge. Mackie's scepticism remains an engaging source of philosophical interest, and while I cannot claim to have overturned his challenge, I hope at least to have shown that the realist need not despair of answering it on its own terms. At the very least, I hope to have shown how several important issues, in both first order ethics and metaethics, can be usefully organized around the central phenomenon of explanation.

Appendices

Appendix A

Normative Explanation and Motivation

A.1 Introduction

In this thesis, I have sought an account of the nature of *normative explanation*, the distinctive normative mode of explanation for overall normative facts. In Chapter 1, I offered two ways of characterising normative explanation. My primary characterisation of normative explanation was through its central rôle in first order normative theory. But, I also followed Korsgaard¹ and Stratton-Lake² in suggesting a secondary characterisation of normative explanation in terms of the motivation of virtuous agents.

Call the idea that there is a correspondence between normative explanation and the motivation of virtuous agents the *Symmetry thesis*. This thesis can be formulated in different ways. Korsgaard presents one formulation, which she attributes to Kant:

The reason why a good-willed person does an action, and the reason why the action is right, are the same.³

¹Korsgaard 1996, p. 60.

²Stratton-Lake 2000.

³Korsgaard 1996, p. 60.

Stratton-Lake argues, however, that this formulation is too strong, being unable to cope with virtuous agents who suffer from false beliefs or occasional *akrasia*. Stratton-Lake thus proposes an alternative formulation:

A good-willed person is disposed to be motivated to do what she ought by the normative reasons why she ought so to act.⁴

Stratton-Lake's objection to Korsgaard rests on an understanding of *having a good will* as a matter of having a medium- to long-term *character trait*. It is arguable, however, that Kant himself was not interested in character traits so much as in *moral worth*, which we might take to something that can be manifested or not on a case by case basis. If we understand the good-willed person simply as someone who is manifesting moral worth, then, we might prefer a third formulation of the Symmetry thesis:

(A.1.1) Agent *s* manifests moral worth iff

1. *s* is motivated to do an act *x*; and
2. *s*' motivation to do act *x*, and the explanation of why act *x* is obligatory, are the same.

In this chapter, I will not try to decide between these various formulations. Instead, I will focus on a claim that they all share in common:

(A.1.2) There are possible cases in which a person manifests a good will in being motivated to perform an action by the very considerations that normatively explain why that action is obligatory.

⁴Stratton-Lake 2000, p. 18.

My aim in this chapter is to defend this minimal version of the Symmetry thesis against some difficulties posed by Jonathan Dancy's recent work on motivation and the explanation of action.

A.2 Dancy on Action Explanation

In much recent work, Jonathan Dancy has defended a sophisticated account of motivation and action explanation.⁵ Dancy's work concerns what I call the *intentional* mode of action explanation. Intentional action explanations are those in which, to borrow John Broome's words, an agent's intentional actions are explained 'in a distinctive way which involves the agent's rational faculty',⁶ more precisely, his capacity for *practical* reason. Intentional action explanations are, for this reason, sometimes called *rationalizing* explanations of action.

To understand Dancy's position, it is helpful to contrast it with the more familiar *Humean theory of motivation*. On the Humean view, intentional actions are to be explained by (a subset of) the agent's mental states of *believing* and *desiring*, with a state of each kind being a necessary part of each such explanation. So if I boil a kettle, pour boiling water through tea leaves, and finally drink the resulting brew, and if, furthermore, I both believe that doing these acts is a way to make a cup of tea and I desire to have a cup of tea, we may have a case where the facts that I have these beliefs and this desire together explain why I acted as I did. Or, if it is

⁵Dancy 1995; Dancy 1996; Dancy 2000; Dancy 2004*c*; Dancy 2006*a*; Dancy 2008.

⁶Broome 2008, p. 55. Broome does not explicitly offer these words as a characterization of a mode of action explanation, so I may be using them against his intentions. Nevertheless, these words serve my purpose well.

not these very beliefs and desires that explain my action, then some other combination of beliefs and desires must do so.

Dancy defends an account of intentional action explanation that is at least two removes from the Humean picture. First, Dancy defends an account of motivation according to which an agent's actions are necessarily to be explained by his beliefs alone, and not by his desires.⁷ Call this view *Pure Cognitivism*.⁸ Second, Dancy's account of intentional action explanation differs from the traditional Humean picture in taking intentional actions necessarily to be explained by an agent's beliefs, not in the sense of the agent's *mental states* of believing, but rather in the sense of the *propositional contents* of those mental states. On Dancy's view, what explains an agent's action is not e.g. that he believes that *p*, but rather the content of his belief, namely that *p*:

It is not our believing that things are so that motivates us [...] but rather what we believe, namely their being so, or that they are so. We have, that is, to look through the believings to the things believed if we are to find the real source of motivation.⁹

Call the view that actions are to be explained by mental states of believing, *Psychologism*. Call Dancy's alternative view, that they are to be explained by the contents of an agent's beliefs, *Anti-Psychologism*.

It will be helpful to think of Anti-Psychologism as the conjunction of two weaker claims. Actions, like perceptions, can be divided into successful

⁷Dancy 1993, Ch. 2; Dancy 2000, Ch. 4.

⁸Pure cognitivism does not imply that desire has no part to play in motivation. On the contrary, Dancy maintains that desire is necessary for motivation, but only because desire *just is* the state of being motivated to act. (Dancy 1993, Ch. 2; Dancy 2000, p. 13)

⁹Dancy 2000, p. 77.

and unsuccessful cases, though different accounts of this distinction are possible. Dancy, as we will see, adopts a minimal conception of success in action, requiring only that the relevant beliefs on which the agent acts are true. However, more demanding notions of success are possible. For example, one might require that the agent's relevant beliefs should be both true and justified, or that they should qualify as knowledge. One might also require that the agent should have reasoned correctly with his beliefs. It is not my task here precisely to delineate the distinction between successful and unsuccessful cases of action. I will assume, simply, that a *necessary* condition of success in action is that all the agent's relevant beliefs are true, and hence, that the falsity of some of those beliefs is a *sufficient* condition of unsuccessful action. Since every case of action is either successful or unsuccessful, Anti-Psychologism is logically equivalent to the conjunction of these two claims:

Success Anti-Psychologism In successful cases, actions are to be explained by the (true) contents of an agent's relevant beliefs.

Failure Anti-Psychologism In unsuccessful cases, actions are to be explained by the (false) contents of an agent's relevant beliefs.

The problems that Dancy's position poses for my formulation of the Symmetry thesis in A.1.2 derive from Failure Anti-Psychologism (hereafter, *Failure AP*). For, on the uncontroversial assumption that there can be unsuccessful cases in which an agent is motivated to act by false beliefs, *Failure AP* entails (as Dancy acknowledges) that action explanation is *non-factive*. For example, if I falsely believe that the hotel is on fire, and this

motivates me to escape by jumping out of the window, *Failure AP* entails that I jump out the window *because the hotel is on fire*, even though it is not true that the hotel is on fire.

There are several reasons to be sceptical of the claim that action explanation is non-factive. I, for one, find it hard to make sense of the claim that a mode of explanation could be non-factive.¹⁰ The non-factiveness of action explanation also leads Dancy to the further controversial conclusion that motivational explanations are not *causal*.¹¹ There are, then, some general reasons to be sceptical of *Failure AP*. But, my reason for examining here is the more specific one that *Failure AP* sits particularly uneasily with my minimal formulation of the Symmetry thesis.

Failure AP poses several subtle difficulties for the connection between normative explanation and motivation. The first difficulty is that, if action explanation is non-factive, we lose what is distinctive and interesting about *Success AP*. For, when *Success AP* is combined with the claim that action explanation is non-factive, it does no justice to the intuitive idea that, when a virtuous agent *successfully* acts for the right reason, say that *p*, he acts because *p* is *true*. This leads, in turn, to two further problems. First, it leaves the successful virtuous agent looking objectionably detached from the considerations he takes to explain his practical choices. For, if action

¹⁰I know of only one other suggestion of a non-factive mode of explanation, namely Hawthorne and Nolan's claim (Hawthorne and Nolan 2006, p. 271) that explanation by Aristotelian final causes, if it existed, would be non-factive. For example, the germination of an acorn might be explained by its final end of becoming a mature oak tree, even if that final end is not realized because the young tree is cut down before maturity. But, the controversial nature of final causation surely derives, at least in part, from its non-factiveness.

¹¹Dancy 2004c, p. 26.

explanation is non-factive, then acting because p does not require the truth of p . Second, if, as I have argued, normative explanation *is* factive, then the virtuous agent also begins to look objectionably detached from the question of whether the considerations that motivate his action *explain* why he ought so to act. If the reason that p is to explain why one has an obligation, p must be true. But, if action explanation is non-factive, so that one can act because p whether or not p is true, then one can act because p whether or not p explains why one ought so to act. In summary, Dancy's non-factive account of action explanation makes it difficult, perhaps even impossible, to conceive of the virtuous agent as someone properly connected to the considerations for which he is acting, and properly connected to the impact of those considerations on the moral status of his actions. Yet, I claim, these connections are essential to the virtuous agent. So if the connection between virtuous agency and normative explanation is to be retained, *Failure AP* must be rejected.

One way to attack *Failure AP* would be to attack *Pure Cognitivism*, which is its necessary condition. I shall not pursue this option because, although my account of normative explanation does not require the truth of Pure Cognitivism, it is preferable, nevertheless, that it should be consistent with it. For the purposes of this chapter, then, I will simply assume Pure Cognitivism. A second way to attack *Failure AP* would be to reject both *Failure AP* and *Success AP*, and hence to reject Anti-Psychologism in its entirety. I suggest, however, that adopting this strategy is to reject one of Dancy's important insights, and hence to throw the baby out with the bathwater. My own strategy, then, will be to offer a tentative defence

of *Success AP*, but to *reject Failure AP*. My position thus amounts to a *disjunctivist* account of action explanation, according to which successful actions are to be explained by the *contents* of an agent's beliefs, while unsuccessful actions are to be explained by the fact that the agent *believes* those contents.

I argue for my position by engaging with Dancy's own arguments. Dancy's work gives a central rôle to *motivating reasons*, so I begin, in § A.3, by clarifying this technical term. In § A.4, I consider three of Dancy's arguments for *Success AP*, two which I take to fail, but one of which I take to be more defensible. In § A.5 I consider and reject Dancy's argument from *Success AP* to *Failure AP*.

A.3 Three Conceptions of *Motivating Reasons*

A central theme in Dancy's recent work is the relation between normative reasons and an agent's *motivating reasons*. However, *motivating reasons* is a technical term with various possible definitions, so I will begin by clarifying three alternative conceptions of it, which I call the *explanatory*, the *mental state*, and the *mental content* conceptions. Clarifying this concept will help to avoid certain errors and confusions.

According to one natural definition, an agent's motivating reasons are the *reasons why he acted*. But, this way of talking is ambiguous, for the reason John Broome explains:

Another confusing feature of English is that the explanation of why a fact obtains is also called ‘the reason’ why it obtains. Here, ‘the reason’ is used in a non-normative sense. In this sense, it may be applied to any explanation, whether of a normative or a non-normative fact. The reason why pigs cannot fly is that they have no wings.¹²

Broome’s point is that, understood in one way, to claim that p is *the reason why* q is simply to claim that p explains why q , perhaps in some particular mode of explanation. Broome’s point explicitly concerns only the singular phrase ‘the reason why’, but it can be extended to the plural phrase ‘the reasons why’. If the explanans of q is a conjunctive proposition, then we can understand ‘the reasons why q ’ to mean the conjuncts of its explanans: if $[p \wedge r]$ explains why q , then p and r are *the reasons why* q , and p and r are each *a reason why* q .

I call the conception of *motivating reasons* as complete action explanations, or conjuncts in complete action explanations, the *explanatory conception*. In its broadest version, the explanatory conception of motivating reasons might involve any mode of explanation, be it causal, inferential, conceptual, or so on. However, I will be concerned only with the intentional mode action explanation, so I will ignore all the others.

A second conception of an agent’s *motivating reasons* takes them to be certain of the mental states of agent. On a traditional, broadly Humean picture, these mental states fall into two classes: beliefs and desires, where both of these are to be understood as the agent’s mental states of believing and desiring, rather than the *propositional contents* of those beliefs or desires. On a more flexible conception, an agent’s motivating reasons might

¹²Broome 2004a, p. 34.

include both other propositional attitudes, such as knowledge, intention, and emotional states, and other non-propositional mental states such as being in pain or being tired.

On the third and final conception, an agent's *motivating reasons* are all and only the *propositional contents* of certain of his mental states, such as the propositional contents of (certain of) his beliefs and desires. I call this the *mental contents* conception of motivating reasons. Following Dancy, I will express the relation between an agent and the propositional contents of the beliefs that were relevant to his action, by saying that an agent *acted in the light of* those contents. According to one possible account, what it is for an agent to ϕ in the light of p is for the agent to employ p as a (perhaps non-redundant) premiss in a piece of practical reasoning that leads, in an appropriate way, to his intentional action of ϕ -ing. Other accounts of this relation are possible, however, and for my purposes I need not adjudicate between them. To say that an agent *acted in the light of* p may seem to imply that p is true. I wish to avoid that implication, because I take it to be uncontroversial that agents can be motivated to act in a way that involves their having false beliefs. To avoid any implication of factiveness, I will take the phrase *acting in the light of* p to be shorthand for the non-factive *acting in the (real or apparent) light of* p .

We have, then, three possible conceptions of motivating reasons: the *explanatory*, *mental states*, and *mental contents* conceptions. In discussing the rôle of motivating reasons in the intentional explanations of action, it is vital to keep these three conceptions apart. That is because the choice of which conception of motivating reasons to work with determines which

issues become matters of definition and which, by contrast, become substantive matters that require *a priori* or empirical support. For example, on the *explanatory* conception, where *motivating reasons* are defined as complete or partial intentional explanations of action, it is true by definition that motivating reasons wholly or partly explain an action, whereas it is a substantive question whether an agent's motivating reasons include any of the agent's mental states or their propositional contents. Of course, even on the *explanatory* conception of motivating reasons, it may be true that an agent's motivating reasons include some of his mental states or their propositional contents; but, if it is true, it is not true by definition.

By contrast, on the *mental state* conception of motivating reasons, it is true by definition that an agent's motivating reasons comprise certain of his mental states, whereas it is a substantive issue whether the agent's motivating reasons ever explain, or partly explain, why he acted. On the *mental state* conception of motivating reasons, it may still be true that an agent's actions are explained by his motivating reasons, that is, by his mental states. But, on the *mental state* conception, that claim, if it is true at all, is a substantive truth. It no longer follows trivially from the definition of a motivating reason.

Lastly, on the *mental contents* conception of motivating reasons, it is true by definition that an agent's motivating reasons are the contents of certain of his mental states, whereas it is a substantive issue whether these propositional contents ever explain an agent's action. For it is not true by definition that the contents of an agent's mental states, rather than, say, the fact that he is in mental states with those contents, explain his behaviour.

Failure to distinguish these three conceptions of *motivating reasons* can lead to certain errors. Consider, for example, John Broome's claim that:

Motivating reasons explain or help to explain what a person does, in a distinctive way that involves the person's rational faculty.¹³

This can be read either as a definition of *motivating reasons* in line with the *explanatory* conception, or as a substantive claim about motivating reasons differently conceived, perhaps as mental states of the agent. Since Broome does not isolate these different conceptions, his claim remains ambiguous. Here the failure to separate alternative conceptions results in ambiguity.

Having distinguished these three conceptions of motivating reasons, we are now in a position to discuss Dancy's recent work. In *Practical Reality*, Dancy explicitly adopts the *mental contents* conception of motivating reasons:

When someone does something, there will (normally) be some considerations in the light of which he acted — the reasons for which he did what he did. There are not so many things that we do for no reason at all. Intentional, deliberate, purposeful action is always done for a reason, even if some actions, such as recrossing one's legs, are not — or not always, anyway. So normally there will be, for each action, the reasons in the light of which the agent did that action, which we can think of as what persuaded him to do it. When we think in terms of reasons in this way, we think of them as motivating. The consideration that motivated the agent was his reason for doing what he did.¹⁴

Dancy's conception of *motivating reasons* involves the genuine or apparent considerations on whose light the agent reasoned his way to action, what I

¹³Broome 2008, p. 55.

¹⁴Dancy 2000, p. 1.

have called the *mental contents* conception. This conception of motivating reasons is perfectly legitimate, but Dancy is not always careful to isolate this conception from others. For example, in chapter 4 of *Practical Reality*, Dancy states that he will argue against the view he calls *Psychologism*, according to which:

[...] our motivating reasons are psychological states of ourselves.¹⁵

Instead, Dancy states his intention to defend the claim that '[m]otivating reasons are what is believed [...]'¹⁶ and that:

[...] no reasons at all, neither motivating nor normative, are psychological states of the agent.¹⁷

Given that Dancy holds the *mental contents* conception of motivating reasons, it is very odd that he should want to defend the claim that motivating reasons are the contents of our mental states rather than those mental states themselves. For on the *mental contents* conception, the claim Dancy intends to defend is true by definition, and the claim he intends to disprove is false by definition. Here the failure to separate alternative conceptions of motivating reasons results in a failure to distinguish trivial from substantive claims about motivating reasons.

In the end, I do not think that Dancy's confusion of these different conceptions of motivating reasons matters much for his argument. Dancy may have misexpressed the conclusions he wishes to argue for, namely the theses

¹⁵Dancy 2000, p. 14.

¹⁶Dancy 2000, p. 101.

¹⁷Dancy 2000, p. 109.

of *Success AP* and *Failure AP*, but his arguments for these conclusions are not so easily dismissed. However, by clarifying these different conceptions of motivating reasons and diagnosing the confusions they lead to early on, it will become easier to assess Dancy's real arguments.

A.4 Three Arguments for *Success AP*

My overall aim in this chapter is to defend *Success AP* but reject *Failure AP*. In defending this position, I assume *Pure Cognitivism*, the claim that actions are to be explained only by reference to an agent's beliefs or their contents, and not by reference to his desires or their contents. I begin by arguing for considering Dancy's own arguments in favour of *Success AP*. *Success AP* claims that, in successful cases, an agent's actions are to be explained by the contents of his relevant, motivating beliefs. Given the assumption of Pure Cognitivism, *Success AP* can be defended by disproving the alternative view, *Success Psychologism*, according to which, in successful cases, actions are to be explained by the fact that an agent *believes* certain contents. In this section I critically assess Dancy's arguments in favour of *Success AP*. I argue that two of these arguments fail, but that a third argument is more defensible.

A.4.1 The Argument from Metaphysical Symmetry

Dancy's first argument against *Success Psychologism* rests on the claim that there is a certain metaphysical symmetry between *motivating reasons* and normative reasons. Dancy discusses this metaphysical symmetry via

two claims, which he calls the *explanatory constraint* and the *normative constraint*:

There is a constraint on any theory about the relation between normative and motivating reasons. This is that the theory show that and how any normative reason is capable of contributing to the explanation of an action that is done for that reason. Call this the ‘explanatory constraint’.¹⁸

[The normative constraint] requires that a motivating reason, that in the light of which one acts, must be the sort of thing that is capable of being among the reasons in favour of so acting; it must, in this sense, be possible to act for a good reason. The explanatory constraint held that all normative reasons should be the right sort of thing to contribute to motivation, since that is what they must be if they are to be capable of contributing to the explanation of action in the right sort of way. The normative constraint goes in the other direction, claiming that motivating reasons should be the right sort of thing to be normative reasons.¹⁹

I shall suppose that normative reasons are provided by *facts*, in the sense of *true propositions*. One peripheral strand of Dancy’s work consists in arguing that normative reasons cannot be provided by facts in this sense. Instead Dancy proposes that both normative and motivating reasons are provided by “things believed”, where these are entities ‘about which we can say comparatively little except that they are neither propositions nor facts.’²⁰ For present purposes, this strand of Dancy’s work can safely be ignored.

¹⁸Dancy 2000, p. 101.

¹⁹Dancy 2000, p. 102.

²⁰Dancy 2000, p. 147.

The argument from metaphysical symmetry relies on a very weak sense of the claim that *it must be possible to act for a good reason*. The metaphysical reading that is being given to this claim is simply that motivating reasons (in a yet-to-be-clarified sense) and normative reasons must belong to one and the same ontological category. This is a substantive claim, but it is weak enough to avoid some problems that beset the stronger reading that I discuss in § A.4.2. Indeed, the claim of metaphysical symmetry between motivating and normative reasons is consistent with its being, after all, impossible to act for the right reason. An evil demon might resolve to kill instantly anyone who was about to act for the right reason, and yet motivating reasons and normative reasons might still belong to the same ontological category.²¹

According to Dancy, the claim of metaphysical symmetry between motivating and normative reasons favours *Anti-Psychologism* over *Psychologism*. *Psychologism*, Dancy argues:

[...] conceives of motivating reasons as psychological states, not as things expressible using that-clauses. The reason why I act

²¹Here is a better example of how metaphysical symmetries can be substantive and interesting but weak enough to avoid common problems. Consider for example the anti-realist claim that all truths are knowable i.e. necessarily for each truth, it is possible that it be known. Straightforward attempts to articulate the intent of this claim immediately fall foul of the paradox of knowability that some truths, such as the proposition that some proposition *p* is an unknown truth, cannot logically be known (see e.g. Williamson 1987). A weaker expression of the anti-realist's intent that avoids this problem is that necessarily all truths are the *right category of thing* to be an object of knowledge. Whether this weaker claim should satisfy all anti-realists is a question I shall not pursue. My own view, however, is that this weaker necessary connection between truth and knowledge is substantive, interesting, and may also be true. It is substantive and interesting because while knowledge may essentially be a propositional attitude, it seems a genuine possibility that truths might divide into e.g. *propositions*, which are, and *propositions**, which are not the right category of thing to be an object of knowledge. It seems a substantive claim that truths do not divide in this way.

will be my belief that p in the sense of my believing that p , not that I believe that p .²²

By contrast, Dancy's own view *Anti-Psychologism*,

[...] has one great advantage over psychologism. This is that it provides as reasons things that can be believed, things properly expressible using that-clauses.²³

The essentials of this argument appear several times in Dancy's recent work.

In an earlier paper, for example, Dancy presents the argument against psychologism thus:

The difficulty that I see in this [viz. psychologism] is that it renders incoherent something that I would be very keen to retain, namely that someone can do an action for the very reason that makes it right: that someone's motivating reason can be a justifying reason. This is ruled out of court because justifying reasons are normative facts, and no normative fact can be a psychological state. The two sorts of reason are metaphysically different, in such a way that nothing could possibly be a reason of both sorts at once.²⁴

This argument is also taken up by Stratton-Lake.²⁵

There is a grain of truth in this argument, but it has nothing like the teeth that Dancy claims for it. *Psychologism* can indeed be roughly expressed as the view that actions are explained by an agent's mental states, and there is indeed a true, but trivial, reading of this claim under which it must be false. Explanantia, I have argued, are propositions, and no mental state is a proposition; so no mental state can explain why an agent

²²Dancy 2000, p. 119.

²³Dancy 2000, p. 121.

²⁴Dancy 1996, p. 172.

²⁵Stratton-Lake 2000, pp. 20-1.

acted. This point is trivial, however, because, strictly speaking Psychologism should be defined as the claim that explanations of action are provided by *facts* about an agent's mental states, such as the fact that the agent believes that p . But, facts about mental states are true propositions, so they are perfectly well expressible using 'that'-clauses. It is misleading, then, to claim that the distinction between Psychologism and Anti-Psychologism can be captured as a distinction between taking the explanantia of intentional action explanations to be mental states and taking them to be propositions. *Every* view about action explanation must hold that their explanantia are propositions. The real distinction between Psychologism and Anti-Psychologism concerns whether the propositional explanantia of intentional action explanations are necessarily facts to the effect that the agent is in certain mental states, or whether they are instead the contents of those mental states,²⁶ and the argument from metaphysical symmetry does nothing to decide between these two possibilities. The argument from metaphysical symmetry therefore fails to support *Anti-Psychologism*.

A.4.2 The Argument from Explanatory Symmetry

The metaphysical symmetry thesis discussed above claims only that motivating reasons and normative reason must belong to the same ontological category. The metaphysical symmetry thesis provides a very weak reading

²⁶I have distinguished between Psychologism and Anti-Psychologism using the distinction between facts about the agent's mental states and the contents of those beliefs. There is one possible kind case in which no such distinction could be drawn, viz. when the agent believes some content, p , where p is the content that the agent believes that p . It is not clear to me whether such recursive contents are possible. In any case, I propose to ignore them here.

reading of the intuitive thought that *it must be possible to act for a good reason*, but stronger readings of this claim are possible, and one of them provides the materials for a better argument from what I call the claim of *explanatory symmetry*. While Dancy does not explicitly advance this argument, something like it seems implicit in his work.

One formulation of what I call the explanatory symmetry thesis is given by Bernard Williams, who writes:

If there are reasons for action, it must be that people sometimes act for those reasons, and if they do, their reasons must figure in some correct explanation of their action (it does not follow that they must figure in all correct explanations of their action.)²⁷

Williams' version of the explanatory symmetry thesis can be understood in various ways. One natural reading is as follows²⁸:

(A.4.1) Necessarily, for all propositions p agents s and acts a , if p normatively explains why agent s ought to perform act a , then it is possible both that p explains why s performs act a and p normatively explains why s ought to perform act a .

A.4.1 expresses a connection between deontic and intentional explanation that parallels similar connections between normative reasons and intentional explanations, and it is vulnerable to the same kinds of counterexample. Consider, for example, Mark Schroeder's case of the surprise party:

Nate loves successful surprise parties thrown in his honor, but can't stand unsuccessful surprise parties. If there is an unus-

²⁷Williams 1981.

²⁸I am grateful to my supervisor for help with this formulation.

pected surprise party waiting for Nate in the living room, then plausibly there is a reason for Nate to go into the living room. There is certainly something that God would put into the ‘pros’ column in listing the pros and cons of Nate’s going into the living room. But, it is simply impossible to motivate Nate to go into the living room for this reason — for as soon as you tell him about it, it will go away. Nate’s case [is a] counterexample to many strong theses about the connection between reasons and motivation.²⁹

Schroeder’s counterexample can be amended to count against A.4.1. Suppose that the fact that there is an unsuspected surprise party waiting for Nate in the living room normatively explains why Nate ought to go into the living room. That fact cannot *also* explain why Nate goes into the living room, because if Nate is motivated by the belief that there is an unsuspected surprise party, the party is, *ipso facto*, not unsuspected. So, if Nate believes that there is an unsuspected surprise party waiting for him, then it is false that there is an unsuspected surprise party waiting for him, and no falsehood can normatively explain an obligation. Thus, A.4.1 must be rejected.

Schroeder’s style of counterexample to A.4.1 can be avoided with the following amendment:

(A.4.2) Necessarily, if there is a proposition p , an agent s and an act a such that p normatively explains why agent s ought to do act a , then it is possible that there is a proposition p' , an agent s' and an act a' such that

1. p' normatively explains why s' ought to do act a' ; and
2. p' intentionally explains why s' does act a' .

²⁹Schroeder 2007, p. 165.

A.4.2 has the form $\Box[p \rightarrow \Diamond q]$, which is equivalent in S4 modal logic to a claim of the form $\Box[\Diamond p \rightarrow \Diamond q]$.³⁰ Furthermore, it can easily be verified that the possibility expressed in the consequent of A.4.2 necessarily implies the possibility expressed in its antecedent. If there is a possible case in which an agent performs an action for the very reason that explains why it is obligatory for him so to act, then there is, of course, a possible case in which something explains why an act is obligatory. A.4.2 is thus logically equivalent to a claim of the form $\Box[\Diamond p \leftrightarrow \Diamond q]$, viz.:

(A.4.3) Necessarily, it is possible that there is a proposition p , an agent s and an act a such that p normatively explains why agent s ought to do act a iff it is possible that there is a proposition p' , an agent s' and an act a' such that

1. p' normatively explains why s' ought to do act a' ; and
2. p' intentionally explains why s' does act a' .

This version of the explanatory symmetry is exceptionally weak. It claims only that there is a necessary connection between the possibility of normative explanation and the possibility of cases in which an agent performs an obligatory action for the very reason that normatively explains why it is obligatory. If, as I suggested § A.1.2, virtuous agency can be understood as performing an action because of the very reason that normatively explains why one ought to perform it, then A.4.3 amounts to the claim that normative explanations are possible iff virtuous agency is possible. It is, of course,

³⁰A formal proof is given in Appendix B.

a starting assumption of my entire project that normative explanations are possible. So, by A.4.3, virtuous action must be possible, too.

A.4.3 offers one intuitive way of clarifying Williams' thought that it must be possible to act for a good reason. But, is it true? The problem with A.4.3, or so it seems to me, is that the intuition behind the thought that it must be possible to act *because of* a good reason might be captured by the weaker view that it must be possible to act *in the light of* a good reason. This would undermine any argument for *Success AP* because it is consistent with an agent's acting in the light of a good reason that his action should nevertheless be explained by his *believing* that good reason to obtain, rather than his action being explained by the good reason directly. The availability of this weaker kind of explanation of virtuous action goes hand in hand with Stratton-Lake's weaker conception of the virtuous agent as someone 'who is disposed to be motivated to do the right thing by psychological states which *have as their content* the normative reason why this act is right [...]'.³¹

Given the availability of this weaker conception of virtuous agency and the weaker way of capturing the intuition behind the thought that it must be possible to act for a good reason, the explanatory symmetry thesis offers no support for *Success AP*, at least not without further supplementation. The same objection, I believe, applies to a modified version of the explanatory symmetry thesis, according to which it must be possible for there to be good *external* motivational reasons. By this I mean the thought that it must be possible for there to be good reasons for action which do not concern the agent's own mental states. As Dancy rightly points out, of the reasons

³¹Stratton-Lake 2000, p. 22.

that ordinary agent's typically offer in support of action, comparatively few concern their own mental states. To borrow some of Dancy's examples, the kind of thing that we typically take to be good reasons include 'that she asked me to do it, that this an opportunity I have long been waiting for, and that I will be too busy to have time to do it next week.'³² However, even if we accept the idea that there must be good external reasons of this kind, and that agents must be able to be motivated by these reasons, this does not obviously support the view that these reasons must directly explain the actions of the agents they motivate. Just as before, we can claim simply that it must be possible that an agent should act *in the light* of these external reasons, but that when this happens, what explains why he so acts is that he believes these reasons to obtain.

To defend *Success AP*, what is needed is some way of bridging the gap between the considerations in whose light an agent acts and the explanation of his action. I turn now to consider an argument that aims to bridge that gap.

A.4.3 The Argument from First Personal Intelligibility

I have now considered two arguments for *Success AP*, and rejected them both. However, Dancy has the resources for a third argument which I find more interesting, and which I will now sketch. I characterized above the intentional mode of action explanation as one in which an agent's action is

³²Dancy 2000, p. 100.

explained ‘in a distinctive way that involves the agent’s rational faculty’. This thought might be further developed as the view that there is a mode of action explanation in which the agent’s own perspective on the world, and his own reasoning about what to do, are given some priority. I take this to be Dancy’s own view of the matter, for he writes:

The aim of the explanation of action is to give, so far as possible, the agent’s own perspective on things, so as to reveal the light in which the action was done.³³

The intentional mode of action explanation, then, is one that gives some priority to the agent’s first personal perspective. Now, if we add the further thought that practical reasoning is reasoning with the *contents* of one’s beliefs,³⁴ then it follows that intentional explanations of action should give some priority to the contents of the agent’s beliefs. It is this claim, I would argue, that makes it possible to bridge the gap between the mere claim that an agent’s is acting *in the light* of *p*, and the claim that the agent is acting *because p*. Intentional explanations of action aim to follow, as far as possible, the agent’s own perspective on his choice of action. And from the agent’s own perspective, the various considerations in the light of which he chooses to act are the ones that explain his decision to act. Put another way, the consideration on whose basis the agent decides to act are the *intentional determinants* of his action. Though Dancy does not make exactly this argument in exactly this way, it is, I think, the key implicit assumption of his recent work.

³³Dancy 2000, p. 108.

³⁴This picture of reasoning is developed in some of John Broome’s recent work. See e.g. Broome 2008.

The above is merely a sketch of an argument, so here is an attempt to put the matter more precisely. Let us suppose that there is such a thing as *practical* reasoning, and that it is analogous to theoretical reasoning in deriving “conclusions” from sets of premisses. I put “conclusion” in scare quotes because there is an old question about whether practical reasoning issues in beliefs, intentions, or even actions. Whatever the answer to that question, it need not concern us here. For simplicity, I will assume that practical reasoning issues in intentions to act. The important point is that the premisses to a piece of practical reasoning are the *propositional contents* of the agent’s relevant beliefs and intentions. Of course, it will typically if not universally be a necessary condition of an agent’s employing these contents as premisses in his practical reasoning that he should believe those contents.³⁵ Nevertheless, the agent reasons with the *contents* of his beliefs, not with the facts that he believes those contents.

Suppose, now, that we have a successful case where the agent engages in practical reasoning from the true premiss that p , and that his practical reasoning issues in an intention to ϕ which he promptly discharges by ϕ – *ing*. Suppose we ask now: why did the agent ϕ ? In particular, what rôle, if any, does the premiss that p play in explaining why the agent ϕ -d? According to the development of intentional action explanation proposed above, the answer to this question ought to give some priority to the agent’s own point of view, and from that point of view, p has a special status. After all, for any given decision to act, an agent will believe many propositions that are

³⁵Cases of make-believe and other forms of pretence may not require belief on the agent’s part.

of no relevance to that decision. The special status of p in the agent's decision to ϕ consists in the fact that, as far as the agent is concerned, he chose to ϕ *because* p was the case. Put another way, the agent takes p to be the *intentional* or *rational* grounds for his action. I claim, then, that those who wish to save the phenomena of first personal practical reasoning should be willing to accept the possibility that there is a mode of explanation in which the premisses that an agent employs in a piece of practical reasons can also, at least in successful cases, explain his subsequent action. The mode of explanation for which this claim holds is the one I call the *intentional* mode.

This completes my argument in favour of *Success AP*. As I mentioned above, nothing in the rest of this thesis depends on *Success AP*, so my defence of it could be excised without significant loss. However, I do take *Success AP* to be a defensible position for the reasons just described, and for the internal coherence of my own project, it would be nice to show how my claims about normative explanation might be consistent with *Success AP* without requiring that I also accept *Failure AP* with its attendant difficulties. My aim in the rest of this chapter, then, is to undermine Dancy's argument from *Success AP* to *Failure AP*.

A.5 Dancy's Argument For *Failure AP*

In § A.4.3 I offered a defence of Dancy's claim that in successful cases, actions are to be explained by the propositions in whose light the agent acted. I called this claim *Success Anti-Psychologism*. In chapter 5 of *Prac-*

tical Reality, Dancy argues in defence of the view I have called *Failure Anti-Psychologism*, the view that, in unsuccessful cases, when an agent acts in the light of beliefs some of whose contents are false, the explanation of why the agent acted as he did is still, at least partially, given by those false contents. In this section I criticise Dancy's argument for this view.

The view that I wish to defend against Dancy is a form of disjunctivism in the explanation of action, according to which successful and unsuccessful cases of action are to be explained differently. Dancy, by contrast, defends a non-disjunctivist account of action explanation that treats both successful and unsuccessful cases alike. Part of Dancy's inspiration derives from Bernard Williams statement that:

The difference between true and false beliefs on the agent's part cannot alter the *form* of the explanation which will be appropriate to his action.³⁶

But, Dancy's defence of *Failure AP* rests on no appeal to authority. It is developed instead from the work of Arthur Collins³⁷ My focus is with Dancy's development of Collins' arguments, but I will occasionally cite Collins directly. Dancy has, I think, two independent lines of argument against the disjunctivism position I would endorse. I argue that both arguments fail, but while the first argument is easily dismissed, I take the failure of the second argument to be of some interest.

³⁶Williams 1981, p. 102.

³⁷Esp. Collins 1997.

A.5.1 The Argument From Altered Motivation

The first of Dancy's arguments against disjunctivism I call the *argument from altered motivation*. I think this argument is very simply and straightforwardly in error, but it is worth discussing it in order to be clear about the details of the disjunctive view that Dancy must argue against.

The version of disjunctivism that Dancy takes as his target is as follows³⁸:

A ϕ s for the reason that *p* iff

1. *either* *p* and that *p* is a good reason for ϕ -ing and *A* ϕ -s in light of *p*
2. *or* it is not the case that *p*, but *A* takes it that that *p* is a good reason for ϕ -ing, and *A* ϕ -s in light of his belief that *p*
3. *or* *p*, but that *p* is not a good reason for ϕ -ing but *A* takes it that that *p* is a good reason for ϕ -ing, and *A* ϕ s in light of the fact that *p*.

The first disjunct represents Dancy's conception of *success* in action, while the second and third disjuncts represent two kinds of *failure*. However, I reject the conception of failure given in Dancy's second disjunct. Recall that the phrase *in the light of* is taken to represent the relation between an agent and the propositional contents employed as premisses in his practical reasoning. Now, in Dancy's first disjunct, which represents the successful case, the agent is being taken to act in the light of *p*. However, on Dancy's version of disjunctivism, the case of failure described in the second disjunct states that the agent acted, not in the light of *p*, but in the light of his belief

³⁸Dancy 2000, p. 140. Numbering added.

that p , which I take to mean, ‘in light of its being (or appearing to be) the case that he himself [the agent] believed that p .’ My reason for dissenting from this disjunctivist account is that, if it were true, the disjunctivist would be constrained to say that in successful cases, the premisses that agents employ in their practical reasoning can straightforwardly concern the external world, but that in at least one kind of unsuccessful case, the agent must instead have reasoned from premisses concerning his own beliefs. But, I see no reason for the disjunctivist to accept this constraint. Indeed, so far as I can see, the disjunctivist has no difficulty in allowing that agents can reason on the basis of false beliefs about the external world. Why should Dancy suppose otherwise?

Dancy considers the view that I propose, writing:

If the reasons that motivate us are not psychological states of ourselves, our believings or belief-states, we might still feel that the things we believe cannot be those reasons either. The main reason for saying so is a worry about the case where things are not as the agent conceives them to be. Surely, in such a case, we cannot say that his reason for acting as he did was that p . We have to say that his reason for acting was that he believed that p .³⁹

But, Dancy argues, genuine cases where an agent acted *for the reason that he believed that p*

[...] are rather unusual, and the way in which they are unusual reveals the falsehood of the new theory as a general view about reasons and motivation. Consider a case where my reason for acting is genuinely that I believe that p . For instance, that I believe that the cliff is crumbling is my reason for avoiding

³⁹Dancy 2000, p. 121.

climbing it, because having that belief I am more likely to fall off (I will get nervous). This is a case where that I believe what I do is genuinely my reason for action, in a way that is independent of whether the belief is actually true. As I might say, whether the cliff actually is crumbling or not doesn't matter. I believe that it is crumbling, and this alone is sufficient to motivate me to stay away from it. I recognize that if the cliff were not crumbling, I would still have just the same reason not to climb it as if it were, so long as I continue to believe it to be crumbling. But this is a quite unusual situation, not at all the normal case.

Dancy's argument rests, I maintain, on a confusion of two senses of *acting for the reason that p*. In one of its senses, this phrase is equivalent to *acting in the light of p*, but this sense of the phrase signifies no more than the relation between an agent and the premisses he employed in his practical reasoning, and is not directly explanatory. The other sense of *acting for the reason that p* indicates instead the distinctive intentional mode of action explanation that we have been discussing throughout this chapter. Now, it is true that for the disjunctivist, cases of failure in action will have to be handled by claiming that the agent acted *because he believed that p*, so we might also say that agent's believing that *p* was his reason for acting. But, here we are using this phrase in its explanatory sense, a sense in which we can also consistently say that the agent still acted *in the light of p*. When an agent decides to act on the basis that *p*, the disjunctivist wishes to say that the agent acted *in the light of p*, but that he acted *because he believed that p*. When disjunctivism is understood in this way, Dancy's charge of unusualness does not apply.

The same confusion between different senses of *acting for a reason* can also be found in Collins. Noting the possibility that a false belief that *p*

might motivate an agent to act, Collins counters:

It is his reason, and in case he is wrong, he acts because he makes this error. No claim asserting some other matter about which he is not in error [e.g. that he believes that p] can be substituted in rendering his reason for acting.⁴⁰

Collins point is correct, but as already discussed, it sits perfectly well with disjunctivism. The disjunctivist claims that agents who act in error act in the light of certain false propositions, and they act because they believe those false propositions. The argument from altered motivation is thus unsuccessful against disjunctivism. Indeed, I take this argument to be uninteresting because its failure rests on a simple ambiguity. I turn now to Dancy's more interesting argument.

A.5.2 The Argument from First Personal Accessibility

In § A.5.1, I considered and rejected Dancy's version of disjunctivism. What confused the issue there was the ambiguity involved in the phrase *acting for the reason that p* . In one sense, this is synonymous with *acting in the light that p* , but in its other, explanatory sense, it is synonymous with *acting because p* , where this is the *because* of intentional action explanation. In presenting my own version of disjunctivism, I will therefore avoid this ambiguous terminology.

My version of action explanation disjunctivism can be presented either as disjunctivism about *acting because p* or as disjunctivism about *acting in*

⁴⁰Collins 1997, p. 123.

the light of p. The second versions yield a more natural presentation, so that is what I will work with.

A ϕ s in the light of p iff

1. *A* believes *p*; and
2. a) *either p* is true, *p* is a good reason for ϕ -ing, *A* is in a good epistemic relation to *p* and employs *p* correctly in his reasoning, and *A* ϕ s because *p*;
b) *or p* is false, and *A* ϕ s because he believes that *p*;
c) *or p* is true, but either *p* is not a good reason for ϕ -ing, or *A* is not in a good epistemic relation to *p*, or *A* does not employ *p* correctly in his reasoning, and *A* ϕ -s because he believes that *p*.

In my disjunctivist account, 3(a) represents the case of successful action, while 3(b)-(c) represent two kinds of failure. As noted before, the bar for success can be placed higher or lower. I have described the successful case as one in which the agent is in a good epistemic relation to *p* and employs *p* correctly in reasoning. These conditions are left intentionally vague and flexible. The key point for our purposes is that the truth of the agent's relevant beliefs is a necessary condition of his success and falsity in those beliefs a sufficient condition of failure.

The core of Dancy's argument against disjunctivism is that it

... makes possible something that is in fact impossible, namely for the agent to explain his action in a way that makes no commitment to the truth of the beliefs that he cites in that explanation.⁴¹

⁴¹Dancy 2000, p. 108.

The impossibility is alleged to arise in cases where the second or third disjuncts of the disjunctivist account of explanation come into play. In such cases, the disjunctivist account presented above takes the agent to act because he believes something, and not because of the contents of his belief. In other words, the second and third disjuncts explain the agent's action by reference to his psychological states, rather than by reference to the contents of those psychological states. But, why should this be thought impossible?

Dancy's argument is sophisticated and subtle, so let me begin by sketching it. Dancy writes:

The simple explanation 'I am doing this because p ' clearly expresses the speaker's endorsement of or commitment to the claim that p . No explanation that obliterated that endorsement would be the correct explanation of the action, since it would fail to give the agent's perspective on things, and hence fail to capture the light in which the action was done.⁴²

As I understand it, this argument relies on our previous characterisation of the intentional mode of action explanation as one in which the agent's own perspective is given some priority. In the agent's mind, there is no difference between the propositions in whose light he decides to act and the rational grounds of his decision to act. The agent therefore believes that he is acting because of the contents of his beliefs. Now, Dancy's point is that in unsuccessful cases of action, the explanation proposed by disjunctivism, namely that the agent is acting because he has certain beliefs, is not one that the agent can consistently accept alongside his existing belief that he

⁴²Dancy 2000, p. 108.

is acting because of the contents of his beliefs. The driving force behind this claim is that since belief is not-factive, ascriptions of belief typically include no commitment to the truth of what is believed. Thus, so the argument has it, the agent cannot believe both that he is acting because p and that he is acting because he believes that p .⁴³

What the above argument seems to show is that the explanation that the agent is acting because he believes that p is not one that is first-personally accessible to him. The final step in Dancy's argument is to claim that, since intentional action explanations give priority to the agent's point of view, the explanation of action *must be* accessible to the agent himself. As Dancy puts it, 'our third-person explanation of the agent's act should be, so far as possible, the agent's explanation [...]'.⁴⁴ But, so Dancy contends, the only explanation that is available to the agent is that he acted because p , and not, as the disjunctivist would have it, that he acted because he believed that p . Therefore, so Dancy concludes, even in unsuccessful cases, the correct intentional explanation of action must be given by the false contents of the agent's beliefs, rather than by the fact that he believed those false contents.

⁴³One potential source of confusion that we can lay aside at this stage is that, despite what was said above, both Dancy and Collins accept that agents very often *do* explain their behaviour by saying such things as 'I'm doing it because I believe that p ', as opposed to the more straightforward 'I'm doing it because p '. The move from 'because p ' to 'because I believe that p ' is termed a 'psychologizing restatement' of the action-explanation, and neither Dancy nor Collins see anything illegitimate in it. One ancillary problem, then, concerns how to understand an agent's psychologizing restatement of the explanation of his action, given that it is not thought possible for it to express the proposition that he is acting because he believes that p . Collins and Dancy offer alternative answers to this question, but common to each is the thought that psychologizing restatements retain all that was contained in the original explanation. In essentials, then, neither Collins nor Dancy see any difference between the two explanations when they offered from the agent's own perspective. See Collins 1997, pp. 113–4; Dancy 2000, pp. 109–10.

⁴⁴Dancy 2000, pp. 109–10.

To show how Dancy's argument is invalid, let us set it out more formally. Our concern is with the kind of case where the second or third disjuncts of a disjunctive account might be thought to be appropriate. A description of such a case that should be acceptable to all runs:

1. p is false
2. Agent A acts in the light of p .
3. *Either* agent A is acting because p *or* agent A is acting because he believes that p .
4. Agent A believes (of himself) that he is acting because p .

Now, according to Dancy, it is also true that:

5. Agent A cannot believe (of himself) both that he is acting because p and that he is acting because he believes that p .

In assessing the truth of (5), it matters greatly what kind of impossibility is at issue. Dancy twice cites⁴⁵ Moore's paradox in support of (5), so I take it that what is being claimed is a *rational* impossibility i.e. the agent is *rationally required* not to believe both that he is acting because p and that he is acting because he believes that p . But, if (5) expresses a rational requirement, then without further support, it is simply incapable of doing the work that the argument against disjunctivism requires of it. In particular, these premisses do not entail:

6. It is not the case that agent A is acting because he believes that p .

⁴⁵Dancy 2000, pp. 108–9.

Indeed, Dancy's premisses do not even entail the weaker claim that the agent is *rationally required* not to believe that he is acting because he believes that *p*. That entailment would require the validity of the inference: *S* believes *p*; it is rationally required that *S* does not believe both *p* and *q*; so, it is rationally required that *S* does not believe *q*. However, many⁴⁶, including Dancy himself⁴⁷, deny that wide-scope rational requirements permit detachability in this way. In this bare form, then, the argument from first personal accessibility is invalid.

For Dancy's argument to be valid, something is needed to bridge the gap between (5) and (6). As I indicated in my sketch, I take the missing premiss to be roughly expressed in the idea that intentional action explanation should give some priority to the agent's own point of view. Or as Dancy puts it, 'our third-person explanation of the agent's act should be, so far as possible, the agent's explanation [...]'.⁴⁸ In § A.4.3 I argued that this line of argument should be accepted in successful cases of action, and it might be thought in consequence thought that consistency would demand that I also accept it in unsuccessful cases of action. However, to argue in this way would be simply to assume the falsity of disjunctivism. The essential feature of a disjunctivist view is that successful and unsuccessful cases of action may be accounted for differently. The charge of inconsistency simply does not apply.

It is one thing for the disjunctivist to avoid the charge of inconsistency, and another to justify the disjunctivist position. So what can be said in

⁴⁶Broome 2004*a*.

⁴⁷Dancy 2000, p. 70.

⁴⁸Dancy 2000, pp. 109–10.

favour of disjunctivism? First, disjunctivism avoids the controversial views to which Dancy's non-disjunctivist position leads him. The disjunctivist need not deny that intentional explanations of action are factive, and need not deny that they are causal. These views are significant costs of a non-disjunctivist position, and both are avoided by accepting the disjunctivist alternative. Second, disjunctivism sits more happily with the accounts of normative explanation and virtuous agency that I have developed in Chapters 1–2. So disjunctivism also avoids the costs of rejecting these accounts. I take these to be good reasons for adopting disjunctivism, but one might well feel that the disjunctivist still has a case to answer in explaining *why*, or better, *how* the same considerations that apply in successful cases of action do not apply in the unsuccessful cases. It is to this final issue that I now turn.

Success AP claims that, in successful cases, actions are to be explained by reference to the contents of the agent's relevant beliefs. In § A.4.3, I defended *Success AP* by arguing that from the agent's own point of view, there was no difference between the premisses in whose light he acted, and which he therefore took as the rational grounds of his decision to act, and the explanation of why he acted. Put another way, I argued that *Success AP* is the only way to save the phenomena of first personal practical reasoning. If disjunctivism still has a case to answer, it is the question of why this same argument does not apply to cases of unsuccessful action. Why should we save the phenomena in the successful case but not in the unsuccessful case? Once this question is properly spelled out, I think the answer is plain: the very point of distinguishing between successful and unsuccessful cases of

action is to distinguish between cases where things really are as they seem and cases where they are not. It is one thing to think that phenomena should be saved in successful cases; it is quite another to think that the phenomena can never be misleading or illusory. On the disjunctivist view, unsuccessful cases of action are simply those in which we are misled by appearances. Put yet another way, there is a certain asymmetry between *Success AP* and *Failure AP*. *Success AP* is the reasonable view that in cases of successful action, things are exactly as they seem. By contrast, *Failure AP* is the very implausible view that, even in cases of unsuccessful action, things can never be other than they seem. The motivation for disjunctivism is that it combines the plausibility of *Success AP* but avoids the implausibility of *Failure AP*.

Let me add some final further support for the disjunctivist view. One way to understand the disjunctivist position is to accept that intentional action explanations give some priority to the agent's point of view, but that there are limits to how far this can be accomplished. Suppose, for example, that agent *A* ϕ s in the light of *p*, and he himself believes that he is ϕ ing because *p*. However, let us suppose that *I* know that *p* is false. Dancy expresses the idea that we should give priority to the agent's point of view as the claim that 'our third-person explanation of the agent's act should be, so far as possible, the agent's explanation'. But, what kind of *possibility* is at stake here? I take it to be rational possibility, but it seems very implausible to claim that in explaining why agent *A* ϕ d, I should be rationally required to put aside my own knowledge that $\neg p$ and reason instead from the false belief that *p*. Of course, this is not to deny that

imagining oneself to be in the agent's shoes can be helpful for all sorts of *other purposes*, for example, assessing the internal rationality of the agent's choice or the moral status of his character. Still, this does not show that, in explaining action, there are no limits to how far the agent's own point of view should be adopted.

As I mentioned above, though I reject Dancy's argument from first personal accessibility, I nevertheless find it an interesting failure. What I find interesting is the possibility that there can be situations in which an agent cannot rationally combine a belief in the true explanation of his action with his existing beliefs. When someone acts because he falsely believes that p , he himself may be rationally unable to assent to that thought. For as Dancy argues, it does indeed look Moore-paradoxical for someone to believe both that he is acting because p and that he is acting because he falsely believes that p . Such rational 'blind spots', as they have come to be called, are a phenomenon of some interest. But, the fact that one cannot rationally incorporate a new belief into one's existing beliefs does not show that belief to be false.

A.6 Conclusion

In this chapter, I have offered a tentative defence of a disjunctivist account of action explanation. I have defended *Success AP*, the claim that, in successful cases, actions are to be explained by the *contents* of the agent's beliefs. But, I have also argued, that in unsuccessful cases, actions are to be explained by the fact that the agent *believes* those contents, and thus, I

have denied *Failure AP*.

This disjunctivist account of action explanation sits well with the idea that there is a necessary connection between normative explanation and the motivation of virtuous agents. In particular, it allows us to claim that there is a tight connection between what motivates the virtuous agent and what explains why the agent ought to be acting as he is *at least in successful cases*. For, in successful cases of virtuous action, the very consideration that explains why the agent *ought* to perform a certain act also explains why he *does* perform that action.

Appendix B

A Formal Proof

My argument in Appendix A relied on the equivalence that $\Box[p \rightarrow \Diamond q] \dashv\vdash \Box[\Diamond p \rightarrow \Diamond q]$. Here is a formal proof in modal logic S4.

I prove first $\Box[p \rightarrow \Diamond q] \vdash \Box[\Diamond p \rightarrow \Diamond q]$.

(1)	$\Box[p \rightarrow \Diamond q]$	(k)	Premiss
(2)	$\neg\Box[\Diamond p \rightarrow \Diamond q]$	(k)	Negated Conclusion
(3)	$\Diamond\neg[\Diamond p \rightarrow \Diamond q]$	(k)	2, MN
(4)	kAl		3, $\Diamond R$
(5)	$\neg[\Diamond p \rightarrow \Diamond q]$	(l)	3, $\Diamond R$
(6)	$\Diamond p$	(l)	5, PTr
(7)	$\neg\Diamond q$	(l)	5, PTr
(8)	lAm		6, $\Diamond R$
(9)	p	(m)	6, $\Diamond R$
(10)	$\Box\neg q$	(l)	7, MN
(11)	$\Box\neg q$	(m)	8, 10, $\Box\Box R$
(12)	$\neg\Diamond q$	(m)	11, MN
(13)	$\Box[p \rightarrow \Diamond q]$	(l)	1, 4, $\Box\Box R$
(14)	$[p \rightarrow \Diamond q]$	(m)	10, $\Box R$
(15)	$\Diamond q$	(m)	9, 14, PTr
	\otimes		12, 15

I now prove the reverse, $\Box[\Diamond p \rightarrow \Diamond q] \vdash \Box[p \rightarrow \Diamond q]$.

(1)	$\Box[\Diamond p \rightarrow \Diamond q]$	(1)	Premiss
(2)	$\neg\Box[p \rightarrow \Diamond q]$	(1)	Negated Conclusion
(3)	$\Diamond\neg[p \rightarrow \Diamond q]$	(1)	2, MN
(4)	kAl		2, $\Diamond R$
(5)	$\neg[p \rightarrow \Diamond q]$	(m)	2, $\Diamond R$
(6)	p	(m)	5, Ptr
(7)	$\neg\Diamond q$	(m)	5, PTr
(8)	$\Box[\Diamond p \rightarrow \Diamond q]$	(m)	1, 4 $\Box\Box R$
(9)	$[\Diamond p \rightarrow \Diamond q]$	(m)	7, $\Box T$
	/ \		
(10)	$\neg\Diamond p$ $\Diamond q$	(m)	8, PTr
(11)	$\Box\neg p$ \otimes	(m)	9, MN; 7, 10
(12)	$\neg p$	(m)	11, $\Box T$
(13)	\otimes		6, 12

Bibliography

- Adams, Robert Merrihew (1979). "Divine Command Metaethics Modified Again". In: *Journal of Religious Ethics* 7. Pp. 66–79.
- Annas, Julia (2008). "Virtue Ethics and the Charge of Egoism". In: *Morality and self-interest*. Ed. by Paul Bloomfield. Oxford: Oxford University Press. Pp. 205–224.
- Aristotle (2000). *Nicomachean ethics*. Trans. by Roger Crisp. Cambridge texts in the history of philosophy. Cambridge: Cambridge University Press.
- Armstrong, D. M. (1978). *Universals and scientific realism*. Cambridge: Cambridge University Press.
- (1983). *What is a law of nature?* Cambridge studies in philosophy. Cambridge: Cambridge University Press.
- Baron, Marcia (1984). "The Alleged Moral Repugnance of Acting from Duty". In: *Journal of Philosophy* 81.4 (Apr. 1984). Pp. 197–220. ISSN: 0022-362X.
- Bird, Alexander (2007). *Nature's Metaphysics : Laws and Properties*. Oxford: Clarendon Press. URL: <http://www.oxfordscholarship.com/oso/public/content/philosophy/9780199227013/toc.html>.

- Blackburn, Simon (1984). *Spreading the word : groundings in the philosophy of language*. Oxford: Clarendon Press.
- (1985). “Supervenience Revisited”. In: *Exercises in Analysis , Hacking, Ian (Ed), 47 67*. Cambridge: Cambridge University Press.
- Bostock, David (1997). *Intermediate logic*. Oxford: Clarendon Press.
- Broad, C. D. (1946). “Some of the Main Problems of Ethics”. In: *Philosophy* 21.79 (July 1946). Pp. 99–117. ISSN: 0031-8191.
- Bromberger, Sylvain (1966). “‘Why’ Questions”. In: *Mind and cosmos : essays in contemporary science and philosophy*. CPS publications in philosophy of science. [Pittsburgh, Pa.]: University of Pittsburgh Press.
- Broome, John (2004a). “Reasons”. In: *Reason and value : themes from the moral philosophy of Joseph Raz*. Ed. by R. Jay Wallace. Oxford: Clarendon Press. Pp. 28–55.
- (2004b). *Weighing lives*. New York; Oxford: Oxford University Press.
- (2008). “Reasoning”. Unpublished draft. Feb. 2008.
- Bykvist, Krister (2003). “Normative Supervenience and Consequentialism”. In: *Utilitas* 15.1. Pp. 27–49.
- Cartwright, Nancy (1983). *How the Laws of Physics Lie*. Oxford: Clarendon Press. URL: <http://www.oxfordscholarship.com/oso/public/content/philosophy/9780198247043/toc.html>.
- Casullo, Albert (September 1984). “Conjunctive properties revisited”. In: *Australasian Journal of Philosophy* 62. 289–291(3). DOI: doi:10.1080/00048408412340073.
- Charles, David (1992). “Supervenience, Composition, and Physicalism”. In: *Reduction, explanation, and realism*. Oxford: Clarendon Press.

- Chisholm, Roderick M. (1963). "Contrary-to-Duty Imperatives and Deontic Logic". In: *Analysis* 24.2. Pp. 33–36. ISSN: 00032638. URL: <http://www.jstor.org/stable/3327064>.
- Collins, Arthur W. (1997). "The Psychological Reality of Reasons". In: *Ratio* 10.2. Pp. 108–123.
- Copp, David, ed. (2006). *The Oxford handbook of ethical theory*. Oxford handbooks in philosophy. Oxford: Oxford University Press.
- Crisp, Roger (2000). "Particularizing Particularism". In: *Moral particularism*. Ed. by Brad Hooker and Margaret Olivia Little. Oxford: Clarendon Press.
- (2006). *Reasons and the good*. Oxford ; New York: Clarendon Press.
- Dancy, Jonathan (1981). "On Moral Properties". In: *Mind* 90.359 (July 1981). Pp. 367–385.
- (1993). *Moral reasons*. Oxford: Blackwell.
- (1995). "The Presidential Address: Why There Is Really No Such Thing As The Theory Of Motivation". In: *Proceedings of the Aristotelian Society* 95. Pp. 1–18.
- (1996). "Real Values in a Humean Context". In: *Ratio* 9.2. Pp. 171–183. URL: <http://dx.doi.org/10.1111/j.1467-9329.1996.tb00103.x>.
- (2000). *Practical Reality*. Oxford: Oxford University Press. URL: <http://www.oxfordscholarship.com/oso/public/content/philosophy/9780199253050/toc.html>.
- (2004a). "Enticing Reasons". In: *Reason and value : themes from the moral philosophy of Joseph Raz*. Ed. by R. Jay Wallace. Oxford: Clarendon Press. Pp. 91–118.

- Dancy, Jonathan (2004*b*). *Ethics without principles*. Oxford: Clarendon Press.
- (2004*c*). “Two Ways of Explaining Actions”. In: *Agency and action*. Ed. by John Hyman and Helen Steward. Royal Institute of Philosophy supplement ; 55. Cambridge: Cambridge University Press.
- (2006*a*). “Acting in the light of appearances”. In: *McDowell and his critics*. Philosophers and their critics. Malden, MA ; Oxford: Blackwell Publishing.
- (2006*b*). “Nonnaturalism”. In: *The Oxford Handbook of Ethical Theory*. Ed. by David Copp. Oxford handbooks in philosophy. Oxford: Oxford University Press. Pp. 122–145.
- (2008). “On How To Act Disjunctively”. In: *Disjunctivism : perception, action, knowledge*. Ed. by Adrian Haddock and Fiona MacPherson. Oxford: Oxford University Press.
- Davidson, Donald (1967). “Causal Relations”. In: *Journal of Philosophy* 64. Pp. 691–703.
- (1969). “True to the Facts”. In: *The Journal of Philosophy* 66.21. Pp. 748–764. ISSN: 0022362X. URL: <http://www.jstor.org/stable/2023778>.
- Depaul, Michael R. (1987). “Supervenience and Moral Dependence”. In: *Philosophical Studies* 51. Pp. 425–439.
- Dreier, James, ed. (2006). *Contemporary debates in moral theory*. Contemporary debates in philosophy. Malden, MA ; Oxford: Blackwell Publishers.

- Feldman, Fred (1978). *Introductory ethics*. Englewood Cliffs, N.J: Prentice-Hall.
- Fine, Kit (2002). "The Varieties of Necessity". In: *Conceivability and possibility*. Ed. by Tamar Gendler and John Hawthorne. Oxford: Clarendon Press. Pp. 253–281.
- Geach, P. T. (1991). "Whatever Happened to Deontic Logic". In: Nijhoff international philosophy series ; v. 41. Dordrecht ; London: Kluwer Academic.
- Hare, R. M. (1986). *The language of morals*. Oxford: Clarendon Press.
- Hawthorne, John and Daniel Nolan (2006). "What would teleological causation be?" In: *Metaphysical essays*. Oxford: Clarendon Press.
- Hempel, C. G. and P. Oppenheim (1948). "Studies in the Logic of Explanation". In: *Philosophy of Science* 15. Pp. 135–175.
- Hooker, Brad (2000). *Ideal code, real world : a rule-consequentialist theory of morality*. Oxford: Clarendon Press.
- Hooker, Brad and Margaret Olivia Little, eds. (2000). *Moral particularism*. Oxford: Clarendon Press.
- Horgan, Terence and Mark Timmons (1992). "Troubles on Moral Twin Earth: Moral Queerness Revived". In: *Synthese* 92.2. Pp. 221–260.
- Hornsby, Jennifer (1997). *Simple mindedness : in defense of naive naturalism in the philosophy of mind*. Cambridge, Mass. ; London: Harvard University Press.
- Jackson, Frank (1998). *From metaphysics to ethics : a defence of conceptual analysis*. John Locke lectures. Oxford: Clarendon Press.

- Jackson, Frank, Philip Pettit, and Michael Smith (2000). “Ethical Particularism and Patterns”. In: *Moral particularism*. Ed. by Brad Hooker and Margaret Olivia Little. Oxford: Clarendon Press.
- Kim, Jaegwon (1984). “Concepts of Supervenience”. In: *Philosophy and Phenomenological Research* 45. Pp. 153–176.
- (1987). ““Strong” and “Global” Supervenience Revisited”. In: *Philosophy and Phenomenological Research* 48.2 (Dec. 1987). Pp. 315–326.
- (1993a). *Supervenience and mind : selected philosophical essays*. Cambridge studies in philosophy. Cambridge: Cambridge University Press.
- (1993b). “Supervenience as a philosophical concept”. In: *Supervenience and mind : selected philosophical essays*. Cambridge studies in philosophy. Cambridge: Cambridge University Press. Pp. 131–160.
- (1993c). “Supervenience for Multiple Domains”. In: *Supervenience and mind : selected philosophical essays*. Cambridge studies in philosophy. Cambridge: Cambridge University Press. Pp. 109–130.
- Korsgaard, Christine M. (1996). *Creating the kingdom of ends*. Cambridge: Cambridge University Press.
- Kraemer, Eric Russert (1977). “Conjunctive Properties and Scientific Realism”. In: *Analysis* 37.2 (Jan. 1977). Pp. 85–86. ISSN: 0003-2638.
- Lance, Mark (2004). “Defeasibility and the Normative Grasp of Context”. In: *Erkenntnis* 61.2-3. Pp. 435–455.
- Lance, Mark and Margaret Olivia Little (2006a). “Defending Moral Particularism”. In: *Contemporary debates in moral theory*. Ed. by James Dreier. Contemporary debates in philosophy. Malden, MA ; Oxford: Blackwell Publishers.

- Lance, Mark and Margaret Olivia Little (2006*b*). “Particularism and Antitheory”. In: *The Oxford handbook of ethical theory*. Ed. by David Copp. Oxford handbooks in philosophy. New York: Oxford University Press.
- (2007). “Where the Laws Are”. In: *Oxford studies in metaethics*. Ed. by Russ Shafer-Landau. Vol. 2. Oxford Studies in Metaethics. Oxford, New York: Oxford University Press.
- Lewis, David K. (1970). “How to Define Theoretical Terms”. In: *The Journal of Philosophy* 67.13. Pp. 427–446. ISSN: 0022362X. URL: <http://www.jstor.org/stable/2023861>.
- (2001). *On the plurality of worlds*. Oxford: Blackwell.
- Little, Margaret Olivia (2000). “Moral Generalities Revisited”. In: *Moral particularism*. Ed. by Brad Hooker and Margaret Olivia Little. Oxford: Clarendon Press.
- Mackie, J. L. (1977). *Ethics : inventing right and wrong*. Harmondsworth: Penguin.
- Majors, Brad (2005). “Moral Discourse and Descriptive Properties”. In: *The Philosophical Quarterly* 55.220 (July 2005). Pp. 475–494.
- McDowell, John (1978). “Are Moral Requirements Hypothetical Imperatives?” In: *Aristotelian Society Supplement* 52. Pp. 13–29.
- McLaughlin, Brian P. (1995). “Varieties of Supervenience”. In: *Supervenience : new essays*. Ed. by Elias E. Savellos and Ümit D. Yalçın. Cambridge: Cambridge University Press.
- Mellor, D. H. (1991). “The singularly affecting facts of causation”. In: *Matters of metaphysics*. Cambridge: Cambridge University Press.

- Moore, G. E. (1912). *Ethics*. Home university library of modern knowledge. London: Thornton Butterworth.
- (1993). *Principia Ethica*. Ed. by Thomas Baldwin. Cambridge: Cambridge University Press.
- Neale, Stephen (2001). *Facing facts*. Oxford: Clarendon Press.
- Nozick, Robert (1981). *Philosophical Explanations*. Cambridge: Harvard University Press.
- Parfit, Derek (1987). *Reasons and persons*. 2nd ed. Oxford: Clarendon Press.
- (2007). “Climbing the Mountain”. Unpublished draft. May 2007.
- (2008). “On What Matters”. Unpublished draft. May 2008.
- Pietroski, Paul M. (1993). “Prima Facie Obligations: Ceteris Paribus Laws in Moral Theory”. In: *Ethics* 103.3. Pp. 489–515.
- Pietroski, Paul M. and Georges Rey (1995). “When Other Things Aren’t Equal: Saving “Ceteris Paribus” Laws from Vacuity”. In: *British Journal for the Philosophy of Science* 46.1. Pp. 81–110.
- Prichard, H. A. (1949). “Does moral philosophy rest on a mistake?” In: *Moral obligation : essays and lectures*. Oxford: Clarendon Press.
- (2002). *Moral writings*. Ed. by Jim MacAdam. British moral philosophers no. 3. Oxford: Clarendon Press. URL: <http://www.oxfordscholarship.com/oso/public/content/philosophy/9780199250196/toc.html>.
- Quine, W. V. (1951). “Main Trends in Recent Philosophy: Two Dogmas of Empiricism”. In: *The Philosophical Review* 60.1. Pp. 20–43. ISSN: 00318108. URL: <http://www.jstor.org/stable/2181906>.

- Rodriguez-Pereyra, Gonzalo (2003). "Review of Neale, Stephen "Facing Facts"". In: *Mind*. 2nd ser. 112.448 (Oct. 2003). Pp. 780–786. ISSN: 1460-2113.
- Roojen, Mark van (1996). "Moral Functionalism and Moral Reductionism". In: *The Philosophical Quarterly* 46.182. Pp. 77–81. ISSN: 00318094. URL: <http://www.jstor.org/stable/2956308>.
- Ross, W. D. (2002). *The right and the good*. Ed. by Philip Stratton-Lake. British moral philosophers. Oxford: Clarendon Press.
- Ruben, David-Hillel (1990). *Explaining explanation*. The problems of philosophy : their past and present. London: Routledge.
- Rupert, Robert D. (2008). "Ceteris Paribus Laws, Component Forces, and the Nature of Special-Science Properties". In: *Noûs* 42.3 (Sept. 2008). Pp. 349–380.
- Salmon, Wesley C. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- (1990). *Four decades of scientific explanation*. Minneapolis: University of Minnesota Press.
- Scanlon, Thomas (1998). *What we owe to each other*. Cambridge, Mass. ; London: Belknap Press of Harvard University Press.
- Schroeder, Mark Andrew (2007). *Slaves of the passions*. Oxford: Oxford University Press. URL: <http://www.oxfordscholarship.com/oso/public/content/philosophy/9780199299508/toc.html>.
- Shafer-Landau, Russ (2003). *Moral realism : a defence*. Oxford: Clarendon Press.
- Sidgwick, Henry (1962). *The methods of ethics*. 7th ed. London: Macmillan.

- Sinnott Armstrong, Walter (2006). *Moral Scepticisms*. New York: Oxford University Press, New York.
- Smith, Michael (1994). *The moral problem*. Philosophical theory. Oxford: Blackwell.
- (2000). “Does the Evaluative Supervene on the Natural?” In: *Well-being and morality : essays in honour of James Griffin*. Ed. by Roger Crisp and Brad Hooker. Oxford: Oxford University Press.
- Sobel, Jordan Howard (2001). “Blackburn’s Problem: On Its not Insignificant Residue”. In: *Philosophy and Phenomenological Research* 62.2 (Mar. 2001). Pp. 361–383. ISSN: 0031-8205.
- Stratton-Lake, Philip (2000). *Kant, duty and moral worth*. Routledge studies in ethics and moral theory. London: Routledge.
- Strawson, P. F. (1985). “Causation and Explanation”. In: *Essays on Davidson : actions and events*. Ed. by Bruce Vermazen and Merrill B. Hintikka. Oxford: Clarendon Press. Pp. 115–135.
- (1990). *Individuals : an essay in descriptive metaphysics*. London: Routledge.
- Streumer, Bart (2008). “Are There Irreducibly Normative Properties?” In: *Australasian Journal of Philosophy* 87.4 (Aug. 2008). Pp. 1–25.
- Taylor, Barry (1976). “States of Affairs”. In: *Truth and meaning : essays in semantics*. Ed. by Gareth Evans and John Henry McDowell. Oxford: Clarendon Press. Pp. 263–284.
- Timmons, Mark (2002). *Moral theory : an introduction*. Elements of philosophy. Lanham, Md.; Oxford: Rowman & Littlefield.

- Tooley, Michael (2003). “Causation and Supervenience”. In: *The Oxford handbook of metaphysics*. Ed. by Michael J. Loux and Dean W. Zimmerman. Oxford: Oxford University Press.
- Vallentyne, Peter (2006). “Against Maximizing Act Consequentialism”. In: *Contemporary debates in moral theory*. Ed. by James Dreier. Contemporary debates in philosophy. Malden, MA ; Oxford: Blackwell Publishing.
- Van Cleve, James (1990). “Supervenience and Closure”. In: *Philosophical Studies* 58.3. Pp. 225–238.
- Van Fraassen, Bas C. (1980). *The scientific image*. Clarendon library of logic and philosophy. Oxford: Clarendon Press.
- Wallace, R. Jay, ed. (2004). *Reason and value : themes from the moral philosophy of Joseph Raz*. Oxford: Clarendon Press.
- Wedgwood, Ralph (2007). *The nature of normativity*. Oxford; New York: Clarendon Press.
- Williams, Bernard Arthur Owen (1981). “Internal and external reasons”. In: *Moral luck : philosophical papers, 1973-1980*. Cambridge: Cambridge University Press. Pp. 101–113.
- Williamson, Timothy (1987). “On the Paradox of Knowability”. In: *Mind* XCVI.382. Pp. 256–261. DOI: 10.1093/mind/XCVI.382.256. URL: <http://mind.oxfordjournals.org>.
- (1999). “Truthmakers and the Converse Barcan Formula”. In: *Dialectica* 53.3-4. Pp. 253–270.
- (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

- Williamson, Timothy (2001). "Ethics, Supervenience and Ramsey Sentences". In: *Philosophy and Phenomenological Research* 62.3 (May 2001). Pp. 625–630. ISSN: 0031-8205.
- (2007). *The philosophy of philosophy*. The Blackwell/Brown lectures in philosophy. Malden, MA: Blackwell Publishing.
- Zangwill, Nick (2006). "Moral Epistemology and the Because Constraint". In: *Contemporary debates in moral theory*. Ed. by James Dreier. Contemporary debates in philosophy. Malden, MA ; Oxford: Blackwell Publishing.
- (2008). "Moral Dependence". In: *Oxford studies in metaethics*. Ed. by Russ Shafer-Landau. Vol. 3. Oxford Studies in Metaethics. Oxford; New York: Oxford University Press.