Introduction

Normativity is an elusive concept, one central to much of ethics and epistemology. Normativity tells us what there is reason for us to do, believe, feel, and so on, and what we ought to do, believe, feel, and so on. In fact, oughts and reasons are not the limit of normativity's purview; permission and interdiction are examples of other important normative concepts.

Normativity is a central concern in ethical theories, with their focus the rightness and wrongness of actions, intentions, desires, and character, among other things. Rightness and wrongness are normative concepts, understandable, perhaps, as moral oughts and interdictions. More broadly in practical philosophy, theories of practical reason are concerned with the normativity of actions and related mental states, as well as the normativity of practical reasoning. Likewise, reasons for belief are an important topic in epistemology, and theoretical reason broadly is concerned with the normativity of beliefs and theoretical reasoning. The study of practical and theoretical reason connects closely with a broad variety of other branches of philosophy, from philosophy of science and philosophy of religion to political philosophy and philosophy of mind.

This thesis focuses on a particular topic and issues relevant to it within the study of normativity. The topic that it examines is normative conflicts. A normative conflict occurs when two or more requirements of normativity are not mutually satisfiable in certain ways. What those ways are, the circumstances in which they might arise, and whether or not the conflicts actually arise or only apparently arise are some of the central

1

concerns investigated here.

Putative normative conflicts are of concern to philosophers working in normativity, because different theories of normativity may allow for or not allow for actual normative conflicts. Differing views about normative conflicts put differing constraints on the shape of a theory of normativity, as the arguments in chapter 3 and chapter 4 of this thesis make apparent.

Before giving an outline of the structure of this thesis, it will be helpful to delimit the thesis's scope by setting out some issues in normativity that will not be addressed. Philosophers who work on normativity frequently speak of oughts and reasons. Oughts and reasons are generally taken to serve as the basic building blocks of any theory of normativity. Some philosophers see ought as the more basic of the two,[1] while others believe that a reason is the primitive term in a normative theory.[2] Others believe that neither can be explained fully in terms of the other.[3] Although elements of the relationship between oughts and reasons will be discussed in this thesis, especially in chapter 4, no comprehensive systematic schematic for that relationship will be presented.

Other important disagreements persist in the literature. There is now a well known debate on the question of whether all reasons (or oughts) are internal reasons (or internal oughts), whether they are all external reasons (or external oughts), or whether they are some mixture of the two. There is another important debate as to whether all normativity is instrumental in nature, or whether normativity also gives us substantive ends. And, of course, there are the debates in metaethics which concern realism versus anti-realism and cognitivism versus non-cognitivism for normative terms.

---

[1] E.g. Broome (2004).
[2] E.g. Scanlon (1998) and Skorupski (1999).

These disagreements are important and interesting, but I have not undertaken study them in on them in this work. I take cognitivism for granted, as I do the view that normativity gives us substantial ends in addition to means. I have two motivations for making these working assumptions and not engaging with the debates mentioned above.

The first motivation for not addressing these topics is that a great deal of excellent work has been done on them already. For example, it is my view that the internalism/externalism question has been settled in favour of externalism about reasons and oughts. The topic is not settled in the literature, of course, but as I am persuaded by the externalists, and the question is far too large to take up in a thesis in which it is not meant to figure prominently, I have simply assumed in most of my discussion that externalism about normative terms is true. By 'externalism' here I mean the view that normative reasons are not, or are not limited to being, mental states. The second motivation is that this thesis is primarily concerned with questions that fall under the scope of what might be called the logic of normativity, although I think the general area might be better thought of in terms of the metaphor of the structure of normativity. For many of these questions little or nothing hinges on whether one is an internalist or externalist about reasons, nor does much hinge on whether the realist or anti-realist is correct. For those parts of my argument where the assumption of externalism affects the soundness of the argument, I would in the spirit of J.L. Mackie ask you to consider the argument as demonstrating what follows if it turns out that externalism is true.[4]

One important substantive assumption that I make throughout the thesis is that cognitivism is true. In some cases, this assumption may not be very important. Alan

---

[3] E.g. Smith (2003), although Smith does not make this point explicitly.
[4] See J.L. Mackie's (1977) apology for the atheistic approach of *Ethics*.

Gibbard[5] and Simon Blackburn[6] have both argued that you can get most of what you want from cognitivism out of a non-cognitivist theory. I do not propose to assess in which cases sophisticated non-cognitivism and cognitivism might differ importantly. As I am committed to cognitivism, I would once again ask the indulgence that these arguments be taken to show what follows for the structure of normativity if cognitivism is true. It will be for a further study to analyse what, if any, are the implications for my arguments should cognitivism be false.

Having said something about what I shall not be addressing, it is time to say what I shall be addressing. This thesis concerns problems in the logic or structure of normativity and, as the title suggests, the central topic is conflicts of normativity. With only these two specifications, there is a wide range of problems that might be covered. What follows is a discussion of what issues are addressed here and why.

My initial intention when I began writing this thesis was to try to exert some order on the concepts of ought and of a reason. In philosophy there are at least three different concepts of a reason, although only one of those concepts, that of a normative reason, will be examined in this work. Philosophy of mind and action employs the concept of a reason for which we believe or act, what Jonathan Dancy calls a 'motivating reason'.[7] A different concept of a reason is used by people who work on theoretical and practical normativity. That is the concept of a normative reason, a reason that counts in favour of something. Finally, there is the general idea of a reason as a component of an explanation, sometimes called an 'explanatory reason'. There are no doubt other concepts of a reason, but these three give a good picture of the general range of

---

[5] Gibbard (1990).
[6] Blackburn (1984) and (1998).

possibilities.

As I was impressed by the work of philosophers in the middle of the last century, who were skilled in producing catalogues of possible uses of the word 'ought' and discussing which were relevant philosophically and which uses could be grouped into related bunches, it seemed a worthy project to go a step further and isolate a number of specific features of normative terms, to explain 'what they mean'. I was quickly outmatched by this task and gave it up, but not before discovering a different strand of normativity research that was more manageable in scope. There are a large number of prominent normative theorists who, while holding wildly differing views on the metaphysics and metaethics of reasons, hold a view that I shall call 'normative separatism'. Normative separatism, having stronger and weaker forms, is the view that the different branches of normativity### practical, theoretical, and any others### are independent of each other. The logical structure of normative separatism is discussed in chapter 4 of this thesis.

Normative separatism is more commonly held than argued for, a fact that contributes to the difficulty of giving a clear explication of it. People with such diverse metaethical perspectives as Peter Railton, Alan Gibbard, John Skorupski, Ralph Wedgwood, and possibly Derek Parfit all nevertheless agree that there are no conflicts between the requirements of practical normativity and the requirements of theoretical normativity. A normative conflict occurs when there are two or more normative requirements, with all of which an agent cannot comply. Sometimes these conflicts, or more accurately putative conflicts, arise when an agent cannot both do what she ought to do and believe what she ought to believe. According to normative separatism, in an

---

[7] Dancy (2000), see the Appendix to chapter 1.

important sense conflicts between theoretical and practical normativity are not conflicts at all; the concept of theoretical normativity is distinct from that of practical normativity, and they are no broader concept of normativity that subsumes both of them. Although we have normativity in the practical sphere and we have normativity in the theoretical sphere, matters in one are never affected by matters in the other.

In this thesis I argue that putative normative conflicts present difficult problems for normative theorists. Accepting that there are real normative conflicts or holding that there are only apparent normative conflicts has important ramifications for a theory of normativity. While I do not argue that a particular account of how to treat putative normative conflicts is correct, I do argue that there is reason to doubt the approach of normative separatists to normative conflicts.

The body of the thesis consists of four chapters. The first two chapters address particular normative claims that figure importantly in the arguments of the final two chapters. In chapter 1, I argue against evidentialism. Evidentialism is the view that all normative reasons for belief are evidential reasons. In chapter 2, I discuss two accounts of rationality, one from Derek Parfit and one from John Broome, and I argue that we need a conception of rationality that is independent of our conception of normativity. The nature of reasons for belief and the normative status of rationality play an important role in the arguments of chapters 3 and 4. Chapter 3 discusses three different types of putative normative conflict, one of which occurs as a result of the possibility of non-evidential reasons for belief, and one of which occurs when rational requirements are given normative force, and one of which occurs when the requirements of both theoretical normativity and practical normativity cannot both be met, In chapter 4, the

ramifications of putative normative conflicts for different possible normative architectures are explored, and a way of accounting for apparent conflicts between differing kinds of reasons for belief is presented.

The argument against evidentialism in chapter 1 begins by claiming that two of the possible supports for evidentialism do not, in fact, support it. The first possible support is the logical structure of reasons for belief. I argue that there is nothing in the logical structure of reasons for belief that tells in favour of evidentialism. The second possible support is doxastic involuntarism, the view that agents do not have direct control over their beliefs. I argue that if doxastic involuntarism tells us anything at all about reasons for belief, what it tells us will be equally problematic for evidentialism as it will be for anti-evidentialism.

After discussing why neither the structure of reasons for belief nor doxastic involuntarism tell in favour of evidentialism, a problem with an account favoured by normative separatists is discussed. There is a commonly held position that says that what appear to be non-evidential reasons for belief are in fact reasons to cause oneself to believe something. I argue that this position has little merit to it. The chapter concludes with a discussion of unstable beliefs and how they present a problem for evidentialism.

The second chapter opens with a discussion of Derek Parfit's account of rationality as given in his paper, 'Rationality and Reasons' in *Exploring Practical Philosophy*.[8] Parfit argues that what it is most rational to desire, or do, is determined by what there would be most reason to desire, or do, were your non-normative beliefs to be true.

Parfit's view is interesting in the present context for two reasons. The first is that,

as I argue, accepting Parfit's view about practical rationality is incompatible with being a separatist about rationality. Rational separatism is the view that practical rationality and theoretical rational are distinct and unrelatable concepts. Rational separatism is parallel to normative separatism, insofar as normative separatism is a claim about the unrelatability of different spheres of normativity and rational separatism is a claim about the unrelatability of different spheres of rationality. Provided that rationality is distinct from normativity, normative separatism need not entail rational separatism and *vice versa*. What is interesting in the case of Parfit's view of rationality is that it may be difficult for him to accept normative separatism without accepting rational separatism.

As there are hints that Derek Parfit is a normative separatist, he may be committed to rational separatism, whether or not he wants to be. While I do not seek to prove that Parfit is a normative or rational separatist, I do argue that if he is a rational separatist, then his theory of what is most rational fails. The discussion of Parfit's theory of rationality shows that a position like his is not compatible with rational separatism. Because in the special case of a theory like Parfit's, rational separatism may be more closely connected with normative separatism than would be the case with other theories, chapter 2 raises some concerns over whether those who accept a Parfit-style view of rationality can also accept normative separatism. The continuing consideration of normative separatism is important because of what normative separatism implies about what I call, in chapter 4, *normative architectures*.

There then follows in chapter 2 a critique of how one might use a particular normative relation, a wide scope ought,[9] as a correct way of representing principles or

---

[8] Parfit (2001).
[9] Broome (2000).

requirements of rationality. I argue that even in the circumstances for which it is the correct normative form of a rational requirement, certain counterexamples suggest that our normative requirements may require us not to follow rational principles. This possibility generates logical space for taking rationality as something both different from normativity and also as something that may not necessarily have normative force in all circumstances. This argument is important in setting up chapter 3, where the relationship between normativity and rationality is central to one of the putative normative conflicts examined there.

Chapter 3 makes good on the title of the thesis and sets out a series of putative conflicts of normativity. I discuss several types of normative conflicts, but distinguish generally between intra-normative and inter-normative conflicts. The former are conflicts that occur within a specific domain of normativity, like the domain of practical normativity or the domain of theoretical normativity. For example, if we accept that there are non-evidential reasons for belief, then we must find some way of adjudicating between evidential and non-evidential reasons for belief when they require us, respectively, to hold and not to hold the same belief. By way of contrast, inter-normative conflicts are those that arise between different domains of normativity, as when doing what we ought to do is not compatible with believing what we ought to believe. This situation can occur when normativity requires us to cause ourselves not to believe what it requires us to believe, for example. Alternatively, we may only have the resources to do what we ought to do or believe what we ought to believe, but not both. I argue that different views on putative normative conflicts may have important implications for the overall structure of a theory of normativity.

The final chapter of the thesis discusses how one might address the various putative normative conflicts. The chapter begins with a discussion of possible normative architectures. Normative architectures are structures that relate types of reasons, e.g. reasons for belief, reasons for actions, reasons for feelings, with sources of normativity, e.g. goodness and evidence. The resources for dealing with putative normative conflicts in each architecture are discussed, and that discussion helps to set up an analysis of a possible way of resolving putative normative conflicts between epistemic and non-epistemic reasons for belief.

Epistemic reasons for belief are facts that stand in an epistemic relation to the belief for which they are a reason. Epistemic reasons for belief include evidential reasons, but also include other sorts of reasons discussed in section 1.5. Non-epistemic reasons for belief are facts that stand in a non-epistemic reasons relation to the beliefs for which they are reasons. A formal method for resolving such conflicts in the form of a defeasing function is given. The defeasing function is subject to multiple interpretations, which are examined at length.

The conclusion of the thesis provides a review of important features of the argument and a discussion of some tasks for normative theorists that are suggested by the conclusions of the arguments herein.

Chapter I: Evidentialism

Introduction

10

Here is a wager.[10] If you believe Augustus's tomb is in Stockholm, then an eccentric millionaire will give you half of his fortune. As it stands, you have recently visited Rome and seen Augustus's tomb there. There is no evidence that suggests the tomb has been moved since your visit; on the other hand, it is not impossible that it has been. Ought you to believe that Augustus's tomb is in Stockholm? If you believe that it is, you will get a large reward. Of course, it is very likely that you will be wrong, but the ill of being wrong on a minor matter seems like a small price to pay for great riches. And there is a small but non-zero chance that you will even be right. It is not impossible that the tomb been moved, just highly unlikely. If you do not believe that Augustus's tomb is now in Stockholm, you will likely be right. Being right about the location of Augustus's tomb, at least under normal circumstances, seems like a poor trade-off for enough money to do as you please.

Yet, the fact that there are many benefits and few harms to believing that the tomb is in Stockholm would not be regarded very broadly among philosophers as a reason to believe that the tomb is in Stockholm. This is because many philosophers[11] think that the only considerations that can count as normative reasons for belief are evidential considerations. The position that the only normative reasons for belief are evidential reasons is called *evidentialism*. This chapter argues that the case for there only being evidential reasons for belief does not look strong and that the onus is on the evidentialist to make the case for her position.

---

[10] Pascal's wager provides a more traditional example. I have avoided using it here, however, as Pascal's wager is flawed; even if we accept all of its premises, it still does not count in favour of religious belief. For a clear account of why, see Hájek (2004).

[11] The list is too long to present exhaustively. A representative sampling of places where this view is espoused or assumed includes Adler (2002), Kelly (2002), Parfit (1984), Railton (1994), Wedgwood (2002*b*).

The question of whether or not there are non-evidential reasons for belief is an important one in its own right. If there are genuine non-evidential reasons for belief, then a complete account of theoretical normativity will have to address them. In the context of this thesis, this chapter's criticism of evidentialism plays an important role.

The primary role that this chapter plays is to lend plausibility to the view that there are non-evidential reasons for belief, or anti-evidentialism. When possible normative conflicts are discussed in chapter 3 and some possible resolutions to them are discussed in chapter 4, non-evidential reasons play an important role in the discussion. By making the case that the burden of proof lies with the evidentialist, the arguments of this chapter aid in making the cases set forth in chapters 3 and 4. This chapter casts doubt on evidentialism in three ways. One is through an analysis of the concept of what reasons are and how they are classified. It is argued that the most intuitive way of classifying reasons does not tell in favour of evidentialism. The second way in which this chapter casts some doubt on evidentialism is to show that the involuntary nature of belief, which is commonly taken as part of the basis for accepting evidentialism, does not support evidentialism. A final criticism of evidentialism is introduced in the form of a thought experiment concerning unstable beliefs.

1.1 Evidential and non-evidential reasons

Reason sentences say that particular multi-place relations hold. I shall treat these relations here as having three places: Fact f is a reason for agent A that he φ.[12] On this

---

[12] It is not hard to imagine expanding the number of places in the relation: '...under circumstances c at time t.…' The questions of whether to add additional relata and which to add are not important to the

view, being a reason is a property of a fact, and it is the property of that fact's standing in a reason relation to an agent and a proposition concerning an action, belief, feeling, or anything else for which there can be a reason.[13] During the remainder of this chapter and this thesis, I shall sometimes discuss reasons to act, believe, feel, and so on, rather than reasons for propositions concerning actions, beliefs, feelings, and so on. I do not mean to be equivocating between actions or belief and action or belief propositions; in all cases it is action or belief propositions that are at issue. However, it is burdensome to use the propositional terminology in the discussion of reasons in all contexts, so sometimes the less precise formulation will be used.

Facts can have the property of being reasons to act, the property of being reasons to believe, or the property of being reasons for anything else (e.g. reasons to admire, fear, desire, etc.), although from here onwards only reasons for beliefs and actions will be mentioned for the sake of economy. To see how the same fact can have the property of being both a reason to believe and a reason to act, consider the following example.

You are on safari in Africa and are in rhino country. While idly lounging in the sun, you suddenly hear a loud rumbling sound characteristic of charging rhinos. The rumbling sound is getting closer and closer. The fact that you hear this sound is a reason for you to believe that there is likely to be an angry rhino charging in your direction. That you hear the sound characteristic of a charging rhino is also a reason to act, namely to dash for shelter in your nearby Landrover. This fact, that you hear the sound

argument here, so they will not be discussed. The account of the logical structure of reasons here is based loosely on positions advocated by Skorupski (forthcoming) and (2000).

[13] It is advantageous to regard the third place in the relation as being occupied by propositions concerning actions, belief, feelings, and so on, as this would allow normal logical operations to be performed on the *relata* that occupy the third place of the reason relation. The alternative would be to regard the third place in the relation as being an

characteristic of a charging rhino, has at least two properties: it is both a reason to believe that there is likely to be an angry rhino charging in your direction and a reason to dash for cover in your safari vehicle. In this example, the same fact is both a reason for belief and a reason for action.

In saying that the fact that you hear the sound characteristic of a charging rhino is a reason for action and a reason for belief, we have assigned to that fact two properties: the properties of being two different *types* of reasons. The notion of a reason type requires some more explanation. To begin with, reason types, i.e. reasons for believing, acting, feeling, and so on, are not distinguished by the particular facts that stand in the particular reason relations. In other words, being a particular type of reason is not an intrinsic property of the fact, but a relational property of it. In the rhino example, the same fact has the property of being two different types of reason- a reason to act and a reason to believe. The type of reason that a fact is, or has the property of being, is given by the sort of thing for which the fact is a reason: beliefs, actions, feelings, and so on. For example, a fact is a reason to act when it stands in the reason relation to an action.

In addition to discussing types of reasons, which distinguish what sort of thing a fact is a reason for, one can discuss the grounds for a reason, i.e. on account of what the reason relation obtains. This consideration is of key importance to this chapter. In the case of belief, one can distinguish between evidential reasons for belief and non-evidential reasons for belief. More will be said about this shortly, but it will be helpful to briefly introduce the evidential/non-evidential reasons distinction now.

Evidential reasons for belief are facts that stand in an evidential relation to the contents of the belief. A non-evidential reason for belief is one in which the fact stands

---

action, belief, feeling, etc., or action-type, belief-type, feeling-type, etc.

in any other reason relation to a belief. The fact that you will be awarded a prize for believing something is a reason for you to believe it, but that fact stands in a non-evidential relation to the belief; that you will win a prize for believing something is not evidence that the belief is true. One cannot distinguish whether a fact is an evidential reason or a non-evidential reason by looking at the fact alone. In the example above, the fact that you hear the sound of a charging rhino is an evidential reason to believe that a rhino is likely to be charging. It is also a non-evidential reason to do something, namely seek shelter in your safari vehicle.

It is important to see that what determines whether a reason is evidential or non-evidential is determined by the nature of the relation between the fact that is the reason and what it is a reason for. A fact is an evidential reason for a belief because it stands in an evidential relation to that belief, and a fact is a non-evidential reason for a belief because it stands in some non-evidential relation to the contents of the belief. The danger in losing track of the relational criteria for identifying whether a reason is evidential or non-evidential is that there may be facts that themselves can loosely be described as being of an evidential or non-evidential nature, but only in the limited sense that it is true that some facts are facts about evidence. For example, it is a fact that your seeing the sunrise is evidence for its being before noon. One might want to call this fact an evidential fact just insofar as it is a fact about evidence. Likewise, it is a fact that if one goes for a walk in the New England winter, then one will be cold without a scarf. One might want to call this fact a non-evidential fact, as it is not about evidence.

Whether or not a fact is evidential or non-evidential in the casual manner suggested above does not determine whether that fact can be or is an evidential or non-

evidential reason. A non-evidential fact, e.g. that when walking in very cold weather, it is better to bundle up than not, can be an evidential reason: it could be an evidential reason for you to believe that you would be better off bundling up while walking in cold weather than not doing so. Likewise, a fact about evidence can be a non-evidential reason. It is a fact that the fact that a rooster has just begun his crowing is evidence that it is before noon. This fact (that the fact that a rooster has just begun his crowing is evidence that it is before noon) is a reason to postpone preparation of supper for several hours when one hears the rooster starting to crow. The first fact stands in an evidential relation to the contents of a belief, while the second stands in a non-evidential relation to an action. That would make the former an evidential reason and the latter a non-evidential one.

So if evidential and non-evidential reasons are genuinely distinct categories, then the distinction cannot be made just by looking at the intrinsic features of the fact, whether it is of the sort of fact that is about evidence or the sort that is not. In arguing this, I have been relying on two sorts of distinction, about which more needs to be said. The first sort of distinction is between evidential and non-evidential reasons; the second is between reasons to believe something and reasons to do something.[14] A fact is a reason to do something (i.e. to act) when it stands in a reason relation to an agent and an action. A fact is a reason to believe something when it stands in a reason relation to an agent and a belief. Thus, fact f is a reason for a belief when it stands in a relation of the form: f is a reason for an agent A to believe x. Fact f is a reason for an action when it stands in a relation of the form: f is a reason for an agent A that he does φ.

---

[14] This distinction assumes that believing is not an action in the narrow sense of the term, or at least is not always an action.

So far two features of reasons have been identified. One feature is the type of reason, determined by what a fact is a reason for. The second feature is that on account of which the reason relation holds. With these two features in hand, we can now make a better attempt at distinguishing evidential from non-evidential reasons. The claim that evidentialists make is that only evidential reasons can count as reasons for belief, where evidential reasons are construed, roughly, as evidence for the contents of the belief. In the context of trying to discover whether or not evidentialists are correct about this matter, it would be question-begging, not to mention uninformative, for evidentialists to define an evidential reason as a reason for belief. But they can distinguish evidential from non-evidential reasons by the difference between what conditions are required for the evidential and non-evidential reason relations to hold. To put it another way, in a reason relation in which fact f is a reason for agent A that he ψ, where ψ can be either an act or a belief, whether the fact is an evidential or non-evidential reason is not determined by any features intrinsic to f or to ψ, but rather by virtue of the relation between f and ψ.[15]

Yet, to spell out just what features of the relation count is not easy. Gilbert Harman, who is not an evidentialist, offers an account of the difference between evidential and non-evidential reasons for belief. It is a perfectly good account of what evidence is:

> R is an [evidential] reason to believe P only if the probability of P given R is greater than the probability of P given not-R.[16]

---

[15] Actions and beliefs are, of course, not propositions. To allow ψ to stand as a place in a relation, we should read 'to ψ' propositionally. Thus in 'The fact that it is raining is a reason for Bob to bring an umbrella', we should parse 'to bring an umbrella' as 'that Bob brings an umbrella'.

[16] Harman (1999), p. 17.

And of non-evidential reasons for belief he writes:

> R is a non[-evidential] reason to believe P if R is a reason to believe P over and above the extent
>
> to which the probability of P given R is greater than the probability of P given not-R.[17]

If read literally, these definitions may not be quite complete as definitions of evidential and non-evidential reasons, however. Consider the case of tautologies. *A priori*, the probability that x is x is 1. I may also have some empirical evidence that x is x. For instance, an expert logician who is always right about logic tells me that x is x is true. So, I now have his testimony, which here counts as a reason for me to believe that x is x. However, the probability that x is x conditional on being told that x is x by a logician is no different from the probability that x is x conditional on its not being the case that I was told that by a logician. A complete definition of evidential reasons would need to account for this sort of case.

Harman's definition of non-evidential reasons for belief looks like it may be right. Taking 'over and above' to modify 'extent', then Harman appears to be saying that any fact that is a reason to believe other than *because of* the probabilistic relationship between the fact and the truth of the belief is a non-evidential reason. Given his definition of evidential reason, this would amount to saying that any reason that is not an evidential reason is a non-evidential reason, a point on which I would concur. In light of this, if we already have a definition of an evidential reason, it would be best to define a non-evidential reason in the very simplest way as a anything that is a reason and stands in

some reason relation other than, or in addition to, an evidential reason relation. This definition allows for the possibility that the same fact will be both an evidential and non-evidential reason for the same belief. That your reliable friend has told you that *b* is the case is an evidential reason for you to believe *b*, as your friend is generally correct. That your friend has told you that *b* is the case is also a non-evidential reason for you to believe *b*, as believing *b* will please your friend.

Although Harman's definition of evidential reasons is incomplete, it certainly points the way towards a more complete definition. Giving a precise definition of an evidential reason for belief may prove to be a rather vexing problem, and it should suffice for current purposes to take the view that a fact is an evidential reason to believe something when it stands in an appropriate evidential relation (i.e. is evidence for the contents of the belief) to the belief's contents. That is to say, a fact is an evidential reason for a belief when that fact is evidence for the contents of that belief.[18] A fact is a non-evidential reason for belief when that fact stands in some other reason relation to the belief (even if it also stands in an evidential relation to the belief). For example, one possible type of non-evidential reason is a prudential reason. A fact is a prudential reason, if it is a reason to believe *x* and it either makes it the case that it is good for the agent that she believes *x,* or it just is that it the fact that it is good for the agent that she believes *x*.

There are other possible notions of evidential reasons besides the notion of a fact

---

[17] *Ibid.*

[18] Skorupski observes in *The Domain of Reasons* that there are some further restrictions that must be placed on what evidence can count as an evidential reason. Skorupski argues that the epistemic accessibility of the evidence to the agent determines whether or not a piece of evidence can be a reason. Because I do not want to take up more detailed questions of the metaphysics of evidential reasons here, I have left the definition more general, recognising that it needs refinement.

that stands in an evidential relation to the contents of a belief. As Peter Railton observes, there are reasons that could be considered evidential in virtue of their role in maximising the difference of 'true minus false' beliefs.[19] He gives the example of its being the case that one could have a reason to believe some false theorem, because believing that false theorem leads to productive work and a commensurate increase in knowledge that would not have occurred without one's having that false belief. This sort of reason, although perhaps characterisable as an evidential one, is not what evidentialists have in mind. They are only concerned with those evidential reasons in which the fact stands in an evidential relation to the belief for which it is a reason.

The question of whether or not something is a reason for belief is a separate one from whether or not a reason is an evidential or non-evidential reason. The most basic available notion of a reason for belief is that of a fact that stands in a reason relation to a belief. It would be a further property of that fact that it stood in an evidential or a non-evidential reason relation to the content of the belief. Because there is a perfectly good criterion for identifying what type of reason a fact is that does not invoke or directly imply that reasons for belief are evidential reasons, the evidentialist must find further arguments to show that reasons for belief are always evidential reasons.

1.2. Does anti-volitionism support evidentialism?

One approach to showing that all reasons for belief must be evidential reasons is to argue that some features of belief are such that only evidence could stand in a reason

---

[19] Railton (1994). Railton cites this type of reason as an example of something that might be wrongly taken to be a reason for belief, but in fact is a reason to be a 'believer of'. It is not clear what the force of

relation to a belief. Among the more important purported differences between belief and action is that belief is commonly thought to be involuntary, at least in normal circumstances, and action is thought to be voluntary. Here I argue that the non-voluntary nature of belief does not tell in favour of evidentialism.

Anti-volitionists believe that under normal circumstances we cannot believe something by an act of the will. In 1.3 there is a brief discussion of two arguments for anti-volitionism and of some limits on arguments for anti-volitionism when anti-volitionism is used as an argument for evidentialism. But we begin by examining how anti-volitionism ties in with evidentialism. A common loose line of reasoning is offered by Christopher Hookway:

> Since beliefs are not actions, and since it is common to deny that belief formation can itself be subject to the will, it is natural to conclude that if such [epistemic] evaluations are the primary focus of evidential evaluations, the systems of norms that guide them are rather different from those that guide our practical reasoning and action.[20]

In other words, because beliefs are involuntary and actions voluntary, we should expect an important difference in the nature of the norms of belief and the norms of action. Evidentialists think this difference is that only evidence can be a reason for belief, whereas non-evidential considerations serve as reasons for action. Thomas Kelly gives voice to the intuition in this way:

> ...The mere realization that my believing some proposition would issue in good consequences

the distinction is.

does not result in my believing that proposition. On the other hand, the realization that I have strong evidence that some proposition is true typically does result in my believing that proposition. With respect to beliefs, practical considerations seem to be psychologically impotent in a way that epistemic considerations are not. And it is tempting to conclude from this that practical considerations are irrelevant to a belief's rationality

Compare the situation with respect to height. One can, of course, make judgements about the expected consequences of being a certain height. For example, I am confident that I am considerably better off, on the whole, being as tall as I actually am as opposed to being two feet shorter. Still, no one would think that it is more rational for me to be some heights rather than others. Moreover, it's plausible to suppose that the *reason why the expected consequences of my being a certain height make no difference to whether or not it is rational for me to be that height derives from my utter lack of control over my height. (Perhaps if I could control my height, then it would be more rational for me to be some heights rather than others.)*[21] [italics added]

Kelly actually brings up two distinct points. One is that there is typically a connection between my judgement that there is evidence for *b* and my believing *b*, while there is no such connection between my judgement that it would be best for me to believe *b* and my believing *b*. The other point, which arises in the second paragraph of the quote, is that this connection has an impact on the rationality of the belief because ought (or reason) implies can### because I cannot voluntarily form beliefs based on judging that it would be good for me to do so. There is a more general issue, one not discussed by Kelly, of whether belief formation is never or nearly never subject to the will or whether belief formation is not subject to the will only or primarily in cases of goodness based

---

[20] Hookway (2000), p. 60.
[21] Kelly (2002), p. 6. Italics added.

judgements.

Other proponents of evidentialism do tackle this last issue and accept a very broad version of anti-volitionism, one that denies any beliefs may be subject to the will. Jonathan Adler devotes an entire chapter of his book to the matter,[22] and more generally in philosophy anti-volitionism is largely, although by no means universally, the norm.[23] Whether or not anti-volitionism is the case is a complex debate in its own right, and one that cannot be settled here. Let us assume for the sake of argument that we cannot will our beliefs, and then we will see what follows. If anti-volitionism is true, then as Hookway suggests, it would not be surprising if there turned out to be an important difference between reasons for belief and reasons for action. We *can* will our actions and some mental states, but we *cannot* will beliefs. This difference in willing *feels* important; yet, I do not think that anti-volitionism about beliefs actually does block the possibility of non-evidential reasons for belief. In fact, I shall argue that the argument from anti-volitionism either shows that there are no reasons for belief at all, not even evidential ones, while there are reasons for action, or that the argument from anti-volitionism fails to tell against the possibility of non-evidential reasons for belief.

One view about normative reasons is that they have to be the sort of reasons that we can make good on,[24] that 'ought' (or 'reason') implies 'can'. Note that it is necessary to employ the awkward expression 'make good on' here, because English lacks a universal verb. If 'do' were a universal verb, then it would be nicer to say that normative

---

[22] Adler (2002), especially chapter 2.

[23] Perhaps the beginning of any modern discussion of anti-volitionism is Bernard Williams's paper 'Deciding to Believe', reprinted in Williams (1973). It has had a broad influence in establishing doxastic involuntarism as the dominant view in current debates. The venerable representative of anti-volitionism is Descartes in his fourth meditation.

[24] I use 'make good on' to mean something like 'act on' and 'believe in virtue of'.

reasons must be reasons for things that we can do. At any rate, that normative reasons

must be the sort of reasons that we can make good on immediately seems to create an

important distinction between reasons for action and reasons for belief. There is a reason

for me to eat a healthy lunch today. When I grasp that reason, I can choose to act on it

(and succeed in acting on it, if it is within my powers to do so), because my actions, or at

least many of my actions, are under my direct control. On the standard picture, the

process for belief works quite differently. Although grasping a reason to believe

something may cause me in some way to believe what there is a reason to believe, it is

not a matter of choosing to believe it. In some cases, recognising that there is a reason

for me to believe something is unlikely to lead directly to my believing it. For example, I

might realise that there is a reason for me to believe that I am 6'0" tall, because I would

have more self-confidence if I did believe that. However, because all the evidence

available to me suggests otherwise, I cannot choose to believe that I am 6'0" tall. I see

that there is a reason for me to believe it, but it is not a reason that I can make good on.

My beliefs are not, at least in general, subject to my direct control. Taking 'ought' to

imply 'can', at least initially, it appears that there will be an important *can* limit on my

beliefs that does not apply to my actions. Seeing that something is good to do, I can

choose to do it. Seeing that something is good to believe, I cannot in general choose to

believe it.

　　Before we can accept this intuition about the implications of 'ought' implies 'can'

for the differences between reasons to believe and reasons to act, we would do well to be

more specific about what sort of possibility is suggested by 'can'. One plausible view for

actions is that it is causal or physical possibility, but not psychological possibility, that is

required. It is never the case that I ought to leap over the Empire State Building, because it is causally impossible. It might be the case that I ought to bungee-jump off of a bridge, even though psychologically I cannot bring myself to do it. On the other hand, there is some *prima facie* plausibility to reading 'can' as psychologically possible. Beliefs are psychological states, so perhaps we should look to an individual's psychology for the possibility limit.

I am not sure what sort of possibility would apply in cases of beliefs. One way of reading anti-volitionism is that we are, so to speak, victims of belief. Beliefs force themselves on us. I see a car in front of me and believe that a car is in front of me, regardless of whether or not I wish to believe that. I can, in other words, psychologically only believe what I end up believing.[25] If the victim view about believing is correct, then we are each in an important respect very much at the mercy of how well our individual cognitive apparatus is set up. Consider a person who, owing to certain limits in her cognitive faculties, cannot understand, and as a result misunderstands, how a particular piece of evidence relates to a belief. She will be stuck with a belief that the evidence does not tell in favour of, a helpless victim of an unwarranted belief. An example brings this situation out more clearly.

Jim is kidnapped by a cult. There he is indoctrinated to believe that scientists are all liars, and that if they say something is the case, that is evidence against its being the case. After ten years in the cult, Jim is brainwashed on this matter beyond recovery. Jim reads in the newspaper that a distinguished zoologist claims to have discovered a new species of mammal### the first such discovery in many years### and this discovery has

---

[25] Of course, even on this picture I do retain some control, by choosing where to direct my attention and what sort of inquiries to pursue. Nevertheless, given those choices, I am simply forced to believe whatever

been scientifically documented with the greatest care by other zoologists. The fact that a distinguished zoologist claims to have discovered a new species of mammal and has documented it carefully is taken by Jim to be a reason for him to believe that no such new species has been discovered. Furthermore, if he were to discover that several peer-reviewed journals had accepted the claim and that other zoologists also had subsequently claimed to have encountered the species, that would be all the more reason in Jim's mind to believe that such an animal has not been discovered. Jim cannot help but believe what he believes in this case; it has become psychologically impossible for him to believe otherwise. Jim is no longer psychologically capable of making good evidential evaluations about matters involving the claims of scientists; in fact he has things backwards and is stuck with being wrong.

In cases of belief, if we take 'ought' to imply 'can psychologically', then it looks like evidentialists have to bite the bullet and say that it is not the case that Jim ought to believe that a new species of mammal has been discovered. This is because Jim cannot make good on the reasons he would seem to have for believing that a new species of mammal has been discovered. The situation is really quite bad for evidentialists, if ought implies psychological possibility in cases of belief, because no matter how poorly our belief acquisition apparatus operates, it will never be the case that we ought not to have the beliefs it psychologically requires us to have. So in the above case with Jim, we cannot say that he ought to believe that a new species has been discovered, or even that he ought not to believe that the odds are that a new species has not been discovered.

If instead of psychological possibility, we retain the standard of possibility normally employed for actions that we ought to perform, then we will be using causal

I end up believing as a result of those actions.

possibility. But taking ought as implying causal possibility does not help the evidentialist at all. It is causally possible to have any belief that can be encoded in your brain. Even if we find it implausible, given Jim's evidential views, that Jim could hold the belief that a new species of mammal has been discovered while also believing that scientists say this claim is true, it is certainly causally possible for Jim to hold both beliefs. In fact, it may not be so implausible to imagine Jim holding both beliefs, as Jim may not have noticed that he held both beliefs, and thus not noticed that they are in conflict. We frequently hold conflicting beliefs because no occasion has arisen where both have been occurrent in us.

An example brings out just how weak a restriction on one's beliefs causal possibility would prove to be. Consider some future neurologist who knows how to rewire people's brains to give them beliefs. The neurologist operates on Jim, giving him the belief that Japan is actually located next to The Congo. He also rewires Jim's brain such that when Jim tries to reason out why he believes that Japan is next to The Congo, he becomes distracted and fails to get anywhere with his introspection. So, even when Jim sees maps showing Japan as an Asian island nation, he still believes that Japan is in Africa. Any effort to explain, rationalise, or justify his belief will not get off the ground with Jim.

Jim neither acquired his belief because of evidence nor does he maintain it because of evidence. Yet, it is causally possible that Jim could have this belief about Japan and maintain it, not because he has or ever had any evidence, but because he is in the unfortunate circumstance of being unable to introspect. If it is causal possibility that counts in terms of what we ought to believe, then it looks plausible that there could be

non-evidential reasons for belief. The example with Jim shows that it is causally possible to have beliefs that are formed and maintained on non-evidential grounds, suggesting that normative reasons for belief cannot be excluded on 'ought' implies 'can' grounds, when we use the 'can' of physical possibility.

It may be tempting to read the 'can' in 'ought' implies 'can' as meaning can will. This is not a plausible interpretation. Doxastic involuntarism excludes willing from the picture *ex hypothesi*, even in cases where one wants to will oneself to believe something on evidential grounds. Physical and psychological possibility are not excluded by doxastic involuntarism, so physical and psychological possibility seem the two important issues to consider under the voluntarist framework.

At this point we should briefly recap. Evidentialists think that anti-volitionism tells in favour of evidentialism. Beliefs are generally responsive to truth considerations and evidence, so we cannot just will ourselves to believe something that is contrary to the evidence. In cases where our cognitive apparatus makes it psychologically impossible for us to understand the evidence or to form a belief based on it, then there cannot be reasons for us to believe that which the evidence suggests is true. If we have to be psychologically capable of believing what we have reason to believe or ought to believe, then two conclusions follow. First, it is the case that we have reason to believe something only if we either have that belief or would have that belief under some psychologically possible circumstance (although then we would only have that reason to believe when in that circumstance). Second, if we cannot psychologically help but have a belief, even if it is acquired through and maintained by poor cognitive practices, then

there are no reasons for us not to have that belief.[26] Many evidentialists would find this situation unacceptable, because they want to be able to say that we ought not to have some of the beliefs that we cannot psychologically shed and that we ought to have certain beliefs that we cannot psychologically acquire. For the evidentialist, anti-volitionism narrowly restricts the domain of reasons for belief to those beliefs that we are psychologically capable of having and maintaining, on evidential or non-evidential grounds.

Alternatively, if the sense of possibility that is required for it to be the case that we can have a certain belief is causal possibility, then it is possible to have beliefs that are neither acquired by nor maintained by any evidence. This leaves open the possibility of non-evidential reasons for belief, because they can be 'made good on' from the perspective of causal possibility just as well as evidential ones. In short, based on how one interprets 'can', there are two possibilities. Either anti-volitionism tells against there being any (or nearly any) reasons for beliefs (other than for the beliefs we have or would have under some psychologically possible circumstances that do not exclude non-evidential grounds for belief), but not in a way that excludes non-evidential reasons in particular, or anti-volitionism places at most a very weak restriction on what reasons there can be for belief, and that restriction does not cut along evidential/non-evidential lines.

1.3. A further problem for anti-volitionism as an argument for evidentialism

---

[26] This point will only be of consequence if one accepts that there are reasons not to believe something.

There is a technical difficulty in employing anti-volitionism in defence of evidentialism. Some arguments for anti-volitionism actually assume evidentialism.[27] Louis Pojman offers two arguments against volitionism, the second of which is his argument from the 'logic of belief'.[28] It is an example of an argument that assumes evidentialism; here it is verbatim:

1. If *A* believes that *p*, *A* believes that *p* is true (analysis of the concept of belief).

2. In standard cases of belief, the truth of *p* is wholly dependent on the state of affairs *s*, which either corresponds to *p* or does not correspond to *p* (and thus makes *p* false).

3. In standard cases of belief, whether or not the appropriate state of affairs *s* that corresponds to *p* obtains is a matter that is independent of *A*'s actions and volitions, but there is some truth connection between *s* and *p* (from 2.).

4. In standard cases of belief, *A* subconsciously believes or presupposes premise 3.

5. Rational believing is defined as believing according to the evidence and fully rational believing as believing simply because of the evidence.

6. Therefore, in standard cases of belief, *A* cannot fully rationally both believe that *p* and that his belief is presently caused by his willing to believe that *p*. Rather, if rational, *A* must believe that

what makes his belief true is a state of affairs *s* which obtains independently of his will.[29]

This argument is generally problematic. There is difficulty in interpreting the first premise, as one can no doubt have beliefs without having the concept of truth, as may be the case with children. A second difficulty is that the argument does not actually tell in favour of anti-volitionism, although Pojman is using it for that purpose. It is an argument in favour of a related but distinct view, that you cannot intentionally believe a proposition and also believe that what caused the belief is an act of the will.[30] This allows for one to believe something as a result of an act of the will but then forget that one believed it for that reason and think that one believed it owing to evidence. Although Pojman seems to recognise that this state of affairs can occur, he still thinks that the argument undermines 'the volitionist thesis that one can acquire beliefs directly by willing to have them'.[31] It does not seem to undermine that thesis. At best it is an argument against the claim that you can believe something and also believe that there is insufficient evidential justification for it.

　　　The worry for the evidentialist is that this sort of argument will count as question begging if he deploys it in support of evidentialism. Premise 5 looks very much like the claim that evidentialism is true. So, if the truth of evidentialism depends on anti-volitionism, this class of arguments for evidentialism is not available, for without premise 5, even a suitably modified argument will fail altogether. This is not a very serious problem for the evidentialist, but it certainly rules out the use of anti-volitionism as a

---

[29] *Ibid.,* pp. 48-49.

[30] Pojman does recognise that there are cases of belief in which the truth of the belief does depend on an act of the will, for example the belief that I will that I go to the store depends on my will. In this argument he is speaking about 'standard' cases of belief, and not this sort.

defence for evidentialism by those who think that we should accept anti-volitionism based on something like this 'logic of belief' argument.

As a very brief digression, it is interesting to note that while there are no doubt grounds, and indeed good ones, for anti-volitionism other than the 'logic of belief' argument, no good additional grounds are provided by Pojman. He does give one other argument, but it does not appear to be very promising. This argument is the argument from the phenomenology of belief, again cited verbatim:

1. Acquiring a belief is a happening in which the world forces itself upon a subject.

2. Happenings in which the world forces itself upon a subject are not things the subject does (i.e. not basic acts) or chooses.

3. Therefore, acquiring a belief is not something a subject does (i.e. is not a basic act) or chooses.[32]

Here I only wish to point out that the first premise seems to assume precisely what is the subject of debate, namely whether beliefs are forced upon us by the world or whether we can will them (i.e. force them on ourselves). This argument appears question-begging, and so we may say that it, too, would be of little use to the evidentialist.

1.4. Against the 'causing yourself' account

---

[31] *Ibid.*, p. 38.
[32] *Ibid.*, p. 40.

Anti-volitionist considerations may have been to some degree influential in the development of another line of evidentialist argument. On this view we can make a distinction between what there is reason for us to believe and what there is reason for us to cause ourselves to believe. The former sort of reason is a purely evidential matter, while the latter admits of all sorts of non-evidential reasons. As Kelly points out,[33] Pascal gives us an interesting example of how anti-volitionism might lead us to that conclusion. Pascal advises us that given our inability to will our beliefs, practical considerations like his wager should guide us to put ourselves in a position where we shall then acquire the right belief. Putting ourselves in a position where we shall acquire a belief is a kind of action, an instance of causing, or at least trying to cause, ourselves to believe something. Because we cannot choose to believe something that is good for us to believe, what there is really reason for us to do, on the cause account, is cause ourselves to believe what it is good for us to believe.

The idea that non-evidential considerations concerning belief are actually reasons to act, in particular to cause ourselves to have a belief, is a commonly held position. Those who think that non-evidential considerations provide us not with reason to believe, but rather with reasons to act, argue against (or at least do not accept) what Derek Parfit calls 'state given reasons' for belief, which are those reasons that come from the desirability of having a certain propositional attitude.[34] State given reasons contrast with 'object given reasons', which are reasons given by the object of the attitude. That a painting is beautiful is an object given reason to admire it; beauty is a feature of the object of my admiring attitude. That admiring the painting will make me happy is a state

---

[33] Kelly (2002), p. 6.

[34] Parfit (2001). For problems with objections to state given reasons, see Rabinowicz and Rønnow-

given reason for me to admire the painting; it is not a feature of the painting that is the reason, but rather an effect of having the attitude. Non-evidential reasons for belief are typically state given reasons. For example, the fact that believing I am popular would make me happy is a reason for me to believe that I am popular. It is not that this fact relates in an evidential way to the contents of the belief for which it is a reason, but rather it is a fact about the good effects having that belief would have on my life. As non-evidential reasons for belief are typically state given reasons, those who do not accept the existence of state given reasons will not acknowledge that there are non-evidential reasons for belief. One way for such people to interpret apparent non-evidential reasons for belief is as non-evidential reasons for causing oneself to believe, where causing oneself to believe is taken to be an action.

Geoff, a wealthy eccentric, has offered Joe half of his fortune if Joe believes next Tuesday that it is Wednesday. There now seems to be an excellent reason for Joe to believe something for which there is no evidential reason. At worst, he will miss a few appointments and endure a little embarrassment for getting the day wrong, but he stands to gain a considerable sum of wealth in exchange. Evidentialists argue that Geoff does not really have a reason to believe that it is Wednesday on the coming Tuesday; he just has a reason to cause himself to believe that it is Wednesday.

There are two arguments that shift the burden to the evidentialists to show why it is that there is only a reason for Joe to cause himself to believe that it is Wednesday on next Tuesday and that there is not also a reason for Joe to believe that it is Wednesday on next Tuesday. The first argument is the weaker of the two. This is the argument from the unity of normativity.

Rasmussen (forthcoming).

Let us assume for the sake of argument that by the evidentialists' lights, there is enough reason for Joe to cause himself to believe that it is Wednesday on next Tuesday that he ought to cause himself to believe that it is Wednesday on next Tuesday. And, let us also assume that there is sufficient evidential reason for Joe to believe that it is not Wednesday on next Tuesday. If we accept that Joe ought on next Tuesday to cause himself to believe that it is Wednesday (because it is good for him to do so), and if we also accept that he ought to believe that it is not Wednesday (because of the evidence), then normativity makes a demand of him that he cannot satisfy. Because causing x implies x, if Joe does what he ought to do and believes what he ought to believe, then he will believe that it is not Wednesday and also believe that it is Wednesday. One might suppose that it is impossible to believe x and also to believe not x. If this supposition is true, then the causing account will place an impossible demand on Joe.

If the supposition is not true, the causing account still places an impossible requirement on Joe, but of a slightly different kind. There seems to be something wrong with an agent who both believes x and also believes not x, that an agent *ought* not both to believe x and believe not x. If it is the case that an agent ought not to have beliefs with contradictory contents, then evidentialism in this case will force the agent to do something that he ought not to do, namely have two beliefs with contradictory content.

This objection has limited bite insofar as the unity of normativity can be denied. If one does not think that practical and theoretical reasons or oughts have a direct relation to each other, then Joe's inability to comply with both his theoretical and practical reasons will not be of much concern. But, for those who accept the unity of normativity, this argument presents a significant problem for evidentialism.

Here is a second objection. Evidentialists and non-evidentialists alike can agree that in the case of Geoff and Joe, whether or not Joe gets the prize depends entirely on what he believes, whether or not he causes himself to believe it; after all, Geoff is awarding the prize not for Joe's causing himself to believe something, but for his believing it. Consider a case with a similar form.

Jill has tickets to a concert. The venue has a rule that no latecomers will be admitted. The fact that no latecomers will be admitted is a reason for Jill to arrive at the concert hall on time. One might also think that there is a reason for Jill to cause herself to arrive on time, because if she arrives on time, she will not miss the concert. Jill can cause herself to arrive on time by leaving for the concert at a suitably early hour, hurrying on her way, etc. As it happens, she can also be caused to arrive on time by someone else, perhaps a friend who will grab her, throw her into a car, drive her there, and carry her to the door before the concert starts.

Regardless of whether she causes herself to arrive on time or whether she is caused to arrive on time by someone else, she gets the prize, admission to the concert, for being there on time, not causing herself to do so. The fact that she will not be admitted late is a reason for Jill to arrive on time. There may also be a reason for her to cause herself to arrive on time, but that reason is dependent on there being a reason for Jill to arrive on time. This is an instance of the normativity of the end, Jill's arriving on time, being transmitted to the means, Jill's causing herself to arrive on time.

In the case of the offer from Geoff to Joe, there, too, is a transmission of normativity from the end to the means. Joe, we suppose, cannot on next Tuesday directly form the belief that it is Wednesday, but causing himself on next Tuesday to believe that

it is Wednesday is a way for him to come to believe it. Without assuming to start with

that there are no non-evidential reasons for belief, it is unclear why we would think that

there is no reason to believe that it is Wednesday but that there is a reason to cause

oneself to believe it, when it looks exactly as if this is an instance where the normativity

of an end, believing it is Wednesday, is transmitted to a means, causing oneself to believe

that it is Wednesday. If the reason attaches to what one gets the prize for in the concert

case, it is unclear, without begging the question against non-evidential reasons for belief,

why the reason should not attach to what Joe receives the prize for in the Wednesday

case. It may be that the two cases are different in virtue of the fact that the Wednesday

case involves a belief and an action, while the concert case involves only actions, but the

burden is on the evidentialist to show why this difference is significant.

Another point bears mentioning here. In some cases an agent cannot believe what

the non-evidential reasons tell her to believe without causing herself to believe it. The

two examples above are such cases. There will be times, however, when an agent can

believe what the non-evidential reasons tell her to believe. There are at least two sorts of

cases where one need not cause oneself to believe something that non-evidential reasons

tell one to believe.

The first sort of case occurs when there is also sufficient evidence for the same

belief. Imagine that believing that you had an appearance in a play on Broadway would

make you happy beyond your wildest dreams. As it happens, you are shown evidence

that you did have an appearance in a play on Broadway, when you were too young to

remember having done so. It seems implausible that in this circumstance there is a

reason for you to cause yourself to believe that you were on Broadway, as it is

unnecessary for you to do so, because you will believe that you were on Broadway as a result of the evidence. Or, if there is a reason for you to cause yourself to believe that you are on Broadway, it is more plausibly because there is a general rule that says you have a reason to take any sufficient means to ends for which you have a reason. If it is because of that rule, then clearly the requirement that you cause yourself to believe something that is good for you is just a special instance of a more general rule that there are reasons for you to take any sufficient means to an end for which you have a reason.

The second sort of case occurs when you already hold the belief for which there is a non-evidential reason. Suppose you believe that there is a lamp in your office. An eccentric millionaire offers you untold riches on the condition that you believe that there is a lamp in your office. There is a non-evidential reason for you to believe that there is a lamp in your office, and you have the belief. Here, it may not even be sensible to speak of your causing yourself to have this belief, as you already have it. And, because causing yourself to have that belief is no longer a means to the end that you intend, there is no more to cause yourself to have that belief to do anything that does not actually cause you not to have the belief.

The causing account must fail if we accept that reason never provides us with two or more all-things-considered oughts that are logically not mutually satisfiable. This is because 'cause' is a success verb, so causing oneself to believe something that one ought not to believe means that one will believe something that one ought not to believe. Any reason that there is for an agent to cause himself to believe something, when there is a non-evidential reason for the agent to believe it, derives from the transmission of normativity from an end for which there is a reason to its means. The onus is on the

38

supporter of the causing yourself account to show that non-evidential reasons for belief are just reasons for causing oneself to believe, rather than that any reasons one has to cause oneself to believe in such cases depend on causing oneself to believer's being a means to an end for which there is reason.

1.5. Unstable beliefs and fixed points: a problem for evidentialism

In this chapter thus far, arguments have been presented that are intended to shift the burden of proof onto the evidentialist to show that there cannot be non-evidential reasons for belief. These arguments have been aimed at showing that some positions that might be taken to support evidentialism do not tell in favour of it. In this section, a different sort of argument against evidentialism is employed. Here I argue that in a relevant case, evidentialism cannot provide reasons to believe### or at least cannot directly provide reasons to believe- what it is correct to believe.

To introduce the problem, an example will be required. Suppose Alice is hooked up to a computer via a telepathy helmet. This helmet accurately reads her beliefs and transmits them to the computer. The computer is attached to a display screen. Alice is then given the following complete set of rules for a game of sorts, which we shall call the numbers game.

When hooked up to the computer, Alice will be told that five seconds from now, a number will appear on the screen, which is presently blank. She is told that the number will be 16, if she forms no belief at all about what number will be on the screen. If, however, she forms any belief about what number will be on the screen, then a different

number will appear, which is half the number that she believes will appear plus 1. So, if she believes the number is 6, it will be 4, and if she believes it will be 10, then it will be 6, and if she believes it will be 5 it will be 3.5 and so on. If Alice forms a new belief in the interval between forming an initial belief and a number's appearing on the screen, then the computer performs the operation again and one half plus 1 of the new number will appear in five seconds. Alice knows all these rules, and faces a problem because of it.

In all cases but one, whatever Alice believes will be wrong. If she believes the number will be 8, it will be 5. Alice knows that if she forms no belief at all about the number, the number will be 16. So, once she recognises that she has no belief about the matter, she will believe that the number is 16, and then of course the number will be 9, and Alice will forever be stuck with unstable beliefs. Alice knows that having any belief but one about the number will make her belief about the number false. This problem will present itself to Alice, unless she believes that the number will be 2, in which case it will be 2, as half of 2 plus 1 is 2.

In the case of the numbers game, the philosophically interesting question is what ought Alice to believe? Because Alice knows that if she has no belief at all, the number will be 16, it appears that it would be wrong for her to ignore that evidence. It is not rational for her to believe that it will be 16, however, because she also knows that if she believes that it will be 16, then it will not be 16. She will face the same problem for all numbers other than 2, and so it seems that one plausible candidate for what she ought to believe is that it will be 2.

As far as I can see, there are three possible reasons for Alice to believe that the

number will be 2. The first reason is that she will be right. The second reason is that she will cease to have unstable beliefs, and the third is that there is evidence that the number will be 2. Only the last of the three reasons is available to the evidentialist, and, as I shall argue, it is the least plausible of the three. The three possibilities will be considered in the order given.

The first claim is that the fact that she will be right is a reason for Alice to believe that the number will be 2. The idea might be something like this. Beliefs aim at truth, and reasons for belief are related to the truth of that belief. Alice's belief about the number will be true if and only if she believes that the number will be 2. Alice knows that she will be right if and only if she believes that it is 2. As her beliefs aim at truth, her beliefs will meet their aim if and only if she believes it will be 2. That she will be right if and only if she believes that it will be 2 is a reason for Alice to believe that it will be 2.

This justification appears to be open to an objection. The objection is that it leads to an overly permissive form of externalism about reasons. That you would be right would then be a reason for you to believe any fact, no matter whether you could have any epistemic contact with it or not. There would be reason to believe facts about epistemically inaccessibly distant parts of the universe, facts about the psychology of people of whose existence you are not aware, and so on.

The objection can be met by providing a strengthened version of this justification. The strengthened version would hold that within the appropriate limits of epistemic accessibility, the fact that you would be right if you believe something is a reason for you to believe it. Because the rules of the numbers game are fully spelled out for her, Alice knows that if she believes that the number will be 2, then it will be 2. A strong limit on

epistemic accessibility would hold that the fact that you would be right is a reason to believe something only if you know that you will be right if you believe it. In the case of the numbers game, there would be a reason for Alice to believe that the number will be 2, because she knows that she will be right if she believes that the number will be 2. The knowledge constraint may be too restrictive, but no doubt some weaker form of the epistemic accessibility constraint on when the fact that you will be right counts as a reason to believe something is available.[35]

This justification is not open to the evidentialist. In order for something to be a reason for a belief, according to evidentialism, it must stand in an evidential relation to the contents of that belief. By evidentialist lights, that she will be right if she believes it is a reason for Alice to believe that the number on the screen will be 2, only if the fact that she will be right if she believes it is evidence that the number will be 2. Using Harman's account of evidence, e is evidence for p when the conditional probability of p given e is greater than the conditional probability of p given not e. A look at Alice's situation shows that if Alice abides by the constraints of evidentialism, the probability that the number on the screen will be 2, given that Alice will be right if she believes it, is no greater than the probability that the number on the screen will be 2, given that it is not the case that Alice will be right if she believes it.

To understand why this is the case, it will be helpful to look first at what it would take for the fact that she will be right if she believes that the number will be 2 to be evidence that the number on the screen will be 2. Let us call the proposition that if Alice believes that the number on the screen will be 2, then it will be 2, p. What it would take

---

[35] John Skorupski provides a sophisticated discussion of different epistemic accessibility constraints on reasons in his forthcoming book, *The Domain of Reasons*.

is for Alice to be more likely to believe that the number on the screen will be 2, if *p* is a fact. That is because Alice's believing that the number on the screen will be 2 is the only circumstance under which the number on the screen would be 2.

On a non-evidentialist account, it is more likely that Alice would believe that the number on the screen will be 2, if *p* is a fact. This is because a non-evidentialist might think, as discussed above, that there is a reason to believe what will be right. Believing that it will be 2 is the only way to be right, and there is a non-evidential reason to be right. If Alice is responsive to her reasons, then she will believe that the number will be 2. On the other hand, evidentialism does not hold that the fact that one will be right if one believes it is a reason to believe something. Thus according to evidentialism, there would be no reason at the start of the game for Alice to believe that the number will be 2. If Alice is responsive only to evidential reasons for believing, there will be no reason to which she can respond that will cause her to believe that the number will be 2.

The importance of this difference between the evidentialist and the non-evidentialist is that if evidentialism is correct, the fact that she will be right if she believes the number will be 2 is not evidence that Alice will believe that the number will be 2. Without there being any evidence that she will believe that the number will be 2, there is no evidence that the number will be 2. So, the probability that the number on the screen will be 2, conditional on its being the case that Alice will be right if she believes it will be 2, is no greater than the probability that the number on the screen will be 2, conditional on its not being the case that Alice will be right if she believes it will be 2. On the other hand, for the non-evidentialist, the probability is greater that the number will be 2 when it is the case that Alice will be right if she believes that the number on screen will be 2 than

it is when it is not the case that she will be right if she believes the number on the screen will be 2. This is because the fact that she will be right if she believes it is a reason for the non-evidentialist, and Alice is responsive to her reasons. Because the evidentialist cannot regard the fact that she will be right if she believes that the number will be 2 as evidence that the number will be 2, that Alice will be right is not a reason for her to believe the fixed point or stable belief.

Although that she will be right is a non-evidential reason for Alice to believe that the number will be 2, it is still an *epistemic reason*. An epistemic reason is a reason that holds because of some non-evidential, but still truth related, considerations. Epistemic reasons might include considerations such as the one just discussed, or they might involve reasons for belief issuing from consistency requirements on one's beliefs.[36] The rejection of evidentialism does not entail accepting that there are non-epistemic reasons for belief, only that there are non-evidential reasons for belief.

The second possible justification for believing the number will be the fixed point (the number which, if believed, would be correct) is that one's beliefs will be unstable otherwise. The reason for Alice to believe that it will be 2 is that otherwise, she will have unstable beliefs in the context of the numbers game. The two most plausible explanations of why there is a reason not to have unstable beliefs are not open to the evidentialist.

The first explanation is that there is something dis-valuable about having unstable beliefs. The disvalue appears to come from pragmatic considerations. The constant adjustment of beliefs as is required in the unstable belief scenario takes up time and

---

[36] I am sceptical that consistency requirements, e.g. that one not hold contradictory beliefs, issue reasons. If they do, however, such reasons would be epistemic but non-evidential.

cognitive resources. Quite aside from any other benefit that might accrue from being right, it would be much more pleasant not to have to adjust constantly one's beliefs about what number it will be. This explanation has some plausibility, and being a pragmatic one, is clearly not open to the evidentialist.

The second explanation of why there is a reason not to have unstable beliefs is that there is a sort of incoherence to having unstable beliefs. Aside from the belief that the number will be 2, any belief one might have about what number will appear on the screen will cause that belief to be false. There is something wrong with having beliefs that cause themselves to be false. Having a belief that you know is false is certainly irrational in an important way. Believing the number will be the fixed point### the number which, if believed, will be the correct number### stops one from having the unstable beliefs which, by definition, have the character of making themselves false. In this instance, the explanation of why there is a reason not to have unstable beliefs is that if one has unstable beliefs, then one will have beliefs that cause themselves to be false, which is irrational.

This explanation is open to the evidentialist. The evidentialist can say that the fact that having a belief will cause that belief to be false is evidence that the belief will be false, and so the evidentialist can say that there is a reason not to have unstable beliefs. However, what the evidentialist cannot do is parlay the reason for not having unstable beliefs into a reason to believe the fixed point.

The evidentialist cannot do this; there is still no evidence that the fixed point is or will be the number on the screen. The only evidence that the content of the fixed point belief is true in this instance would be that Alice believes, or is likely to believe, that the

number will be 2. However, the fact that there is a reason for Alice not to believe the unstable beliefs is not evidence that Alice will believe or is likely to believe the stable one. In the numbers game, if Alice is responsive to her reasons, and, as argued above, if there are only evidential reasons, then Alice will simply not believe any of the unstable beliefs. The fact that she will not believe any of the unstable beliefs is not evidence that she will believe the stable one. Without some sort of evidence that she will believe the stable belief, there is no evidential reason for Alice to believe that the number will be 2.

Evidentialists cannot claim that there is a reason to believe the number will be the fixed point in the numbers game, because one will be right if one believes it. They also cannot claim that there is a reason to believe the number will be the fixed point because otherwise one's beliefs will be unstable. What evidentialists need as evidence is for it to be likely that Alice will believe that the number will be 2. The fact that Alice is likely to believe that the number will be 2 is evidence that the number will be 2.

To show that Alice is likely to believe that the number will be 2, the evidentialist could use the technique employed above, using Alice's responsiveness to her reasons as evidence that she will believe that the number will be 2. Earlier in the chapter,[37] the idea that non-evidential reasons to believe are not really reasons to believe, but rather are reasons to cause yourself to believe, was introduced. On this view, any of the non-evidential reasons to believe discussed above could be taken as a reason for Alice to cause herself to believe the fixed point. Alice is responsive to her reasons, and so she is likely to cause herself to believe that the number will be 2. Causing herself to believe that the number will be 2 will result in her believing that the number will be 2, so Alice is likely to believe that the number will be 2. That Alice is likely to believe that the number

will be 2 is evidence that the number will be 2. So, there appears to be an evidential

reason for Alice to believe that the number will be 2.

There are two difficulties with this solution to the numbers game. The first

difficulty is one that I have already discussed,[38] which is that it falls to the evidentialist to

show that there is a reason to cause yourself to believe something when there appears to

be a non-evidential reason to believe it. It is not entirely clear that evidentialism can help

itself to the assumption that putative non-evidential reasons for belief are, in fact, reasons

for causing oneself to believe.

The second difficulty reinforces the first. By a slight modification of the rules of

the numbers game, it is possible to make causing yourself to believe something always

result in your having an unstable belief. In the modified numbers game, the computer to

which Alice is attached is able to detect whether or not she caused herself to have the

belief she has about the number. If she causes herself to have the belief about the

number, then the normal rule of taking her belief, dividing it by 2, and adding 1 no longer

applies. Instead, the computer will display the number that Alice caused herself to

believe multiplied by 2 with 1 added to the product. So, if Alice causes herself to believe

that the number is 2, then the number displayed will be 5. Causing herself to believe

something will ensure that Alice's beliefs are unstable.

Once the new rule has been added and Alice's beliefs will be unstable if she has

caused herself to believe them, there is no longer any evidence that Alice will believe that

the number will be 2, based on her being responsive to reasons. This is because there is

no longer a reason for Alice to cause herself to believe the fixed point, as doing so will

---

[37] See section 1.4, The 'causing yourself' account.
[38] See section 1.4.

ensure that her beliefs are unstable. There is no reason for Alice to think that she will

cause herself to believe the fixed point, and thus there is no evidence that she will believe

the fixed point. The evidentialist cannot provide a reason for believing the fixed point in

the modified numbers game, even under the generous assumption that he might be able to

in the original numbers game. On the other hand, the position of the non-evidentialist is

unaffected by the modification to the numbers game, as there are good non-evidential

reasons for Alice to believe, not to cause herself to believe, that the number will be 2.

The numbers game represents a difficult problem for the evidentialist. There are

at least two plausible considerations in favour of believing that the number will be the

fixed point, the fact that doing so is the only way to be right and the fact that doing so is

the only way to avoid having unstable beliefs. Neither of these is available to the

evidentialist. These considerations in favour of believing the number will be the fixed

point intuitively look like good reasons for believing the number will be the fixed point,

and evidentialism's inability to accommodate them is problematic. The modified version

of the numbers game shows that even if there is some initial plausibility to the view that

there is a reason to cause yourself to believe something when there appears to be a non-

evidential reason to believe it, there will be situations in which there is clearly no reason

to cause yourself to believe something on those grounds.

Conclusion

In this chapter I have argued that evidentialism has a difficult time getting off the

ground. Doxastic involuntarism, conceptual or contingent, does not provide support for

evidentialism, nor is there anything in the logical structure of reasons for belief that *prima facie* suggests that there cannot be non-evidential reasons for belief. Neither of these arguments show that evidentialism is false, but they both suggest that support for evidentialism will have to come from a different line of argument.

The numbers game represents a very different challenge for evidentialism, and its implications for evidentialism are open to interpretation. The numbers game does not question the justification for evidentialism, but rather raises doubts as to whether or not evidentialism could be true at all. The numbers game has three possible implications for evidentialism.

The first possible implication is that evidentialism, at least in its strict form, is false. This implication would be born out in the following form. There is a reason to believe that the number will be the fixed point. Evidentialism cannot provide a reason to believe the number will be the fixed point. So, evidentialism is false.

The second possible implication is that evidentialism can only be viable as a theory of reasons for belief in certain circumstances. Cases in which there is no evidence that the number will be the fixed point require a separate theory. On this view, evidentialism is shown not to apply to cases like the numbers game, while it is still the correct theory of theoretical reasons for more standard beliefs.

The final possible implication of the numbers game is that while it may be intuitive to believe that there is a reason to believe that the number will be fixed point in the numbers game, there actually is no reason to believe it, because there is no evidential reason for the fixed point. This reading of the numbers game has the least plausibility, because it implies that theoretical normativity is silent on epistemic situations of a

parallel form to the numbers game.

I believe that the first reading of the numbers game has the greatest plausibility. The evidentialist cannot deliver a reason to believe that the number will be the fixed point, when it is clear that there is a reason to believe that the number will be fixed point. There is little plausibility to the third reading, which leaves the second one as a middle view. In order for the second view to be plausible, the evidentialist would have to provide an explanation of why evidentialism is a true theory of reasons for belief for some types of belief but not for others. Once again the burden is shifted to the evidentialist to show why evidentialism is a plausible view.

While this chapter has not proved evidentialism to be false, it has pushed the burden of proof to the evidentialist, leaving open the possibility that there are non-evidential reasons for belief. Doing this is important in the context of the thesis, because in chapters 3 and 4 the arguments frequently invoke non-evidential reasons for belief. Because evidentialism is a view that often lurks as a background assumption in the discussion of reasons for belief, it is important to show that non-evidentialism has an equal, if not greater, plausibility if non-evidential reasons for belief are to be invoked in the arguments of this thesis.

Chapter II: Normativity and Rationality

Introduction

In this chapter I shall examine the relationship between normativity and rationality and shall argue that there are grounds for thinking that rationality cannot be reduced to a sub-species of normativity. The chapter will first focus on a distinction that Derek Parfit makes between what there is most reason to do and what it is most rational to do. His claim is that it is most rational to do what one would have most reason to do, if one's relevant non-normative beliefs were to be true. I shall argue that Parfit's account is problematic and that explicating the notion of what is most rational in terms of what there is most reason for does not look like a successful approach. I shall also discuss the work of John Broome. Broome suggests that one might account for the difference between what one has most reason to do (or what one ought to do) and what it is most rational to do as being a difference in the scope of the normativity involved. I shall argue that while this approach is more promising, it is ultimately subject to problems that point towards general difficulties in explicating rationality in normative terms.

This chapter serves three purposes within the thesis. The first is to bring into clearer focus the relationship between normativity and rationality. While this project is valuable for its own sake, having some understanding of this relationship is important to understanding the structure of the normative conflicts between the normative requirements of practical rationality and those deriving from goodness that are set out in

chapter 3. In chapter 3 it is argued that rational requirements may entail some sort of

normative requirement. The analysis in this chapter shows that while rational

requirements may entail normative requirements, rational requirements themselves are

not just a special class of normative requirements. The second purpose of this chapter is

to continue the sub-thread that runs through this thesis that casts doubts on normative

separatism. In critiquing Derek Parfit's views, problems with the compatibility of

normative separatism with Parfit-style theories of rationality are shown. Finally, this

chapter serves to introduce important concepts concerning the scope of normative and

rational operators that will be relied upon in chapters 3 and 4. The chapter begins with a

critique of Parfit's account of rationality.

2.1. Parfit's account

In his paper 'Rationality and Reasons', Parfit distinguishes between what we

have most reason to want and do and what it is most rational for us to want and do.[39] For

Parfit, the questions of what it is we have most reason to want or what it is most rational

to want and what we have most reason to do or what it is most rational to do are closely,

in fact inseparably, linked. For purposes of this study, I shall not consider the question of

what we have most reason to want or what it is most rational for us to want. Instead, the

focus only will be on the doing. As far as the analysis here is concerned, there is not

anything significant at stake in leaving out the wanting, insofar as Parfit thinks that

reasons for wanting and reasons for doing are the same.[40]

---

[39] Parfit (2001).

[40] Parfit holds this view, at least in part, because he discounts the possibility of there

Parfit sets out the distinction between what we have most reason to do and what it is most rational to do in the following way:

> ...While reasons are provided by the facts, the rationality of our desires and acts depends instead on our beliefs. When we know the relevant facts, these questions have the same answers. But if we are ignorant, or have false beliefs, it can be rational to want, or do, what we have no reason to want, or do. Thus, if I believe falsely that my hotel is on fire, it may be rational or me to jump into the canal. But I have no reason to jump. I merely think I do. And, if some dangerous treatment would save your life, but you don't know that fact, it would be irrational for you to take this treatment, but that is what you have most reason to do....[41]

Parfit's assertion is that what you have most reason to do depends on the facts, while what it is most rational for you to do depends on your beliefs. This philosophical distinction mirrors a common one. In ordinary usage, the distinction typically occurs when we assert that somebody's behaviour 'makes sense' or is rational in light of what she believes, but that she still did not do what she ought to have done, because her beliefs were false.

Although Parfit has made a philosophical distinction, one with a good intuitive basis, of some importance, it is not clear that Parfit has made it in quite the right way. To begin with, we need an analysis of what it is for us to have most reason to do something. Parfit means by this locution that we ought to do it. There are other possible uses of 'most reason'; it is not *a priori* the case that one ought to do that which there is most

---

being what he calls state-given reasons for belief. Chapter 1 of the thesis argues for the possibility that there may be state given reasons for belief. The most comprehensive current paper on the matter is Rønnow-Rasmussen and Rabinowicz's 'Strike of the Demon' forthcoming in *Ethics*.

reason for one to do, at least if most reason is taken to imply that reasons are weighed up

and that one ought to do that for which there is the greatest weight of reason. However,

not all normative theories are committed exclusively to weighing.[42] A theory could, in

principle, have two separate basic normative concepts, *ought* and *reason*. In such a

theory, reasons are weighed against each other, but there are some oughts that are

absolute and admit of no weighing. Perhaps such *oughts* are basic requirements or are

derived by some sort of consistency and universalisability requirements, as Kant

imagined was the case for lying. Such a theory might then hold that other considerations,

for which there are no consistency and universalisability arguments, are best dealt with

through a weighing approach. According to such a theory, when lying is involved, there

is no weighing to be done; we simply ought not to lie. However, when there is a choice

between saving the family dog and the family cat from a fire, then we weigh up the

reasons and in that instance do what we have most reason to do. Whether or not Parfit is

committed to a strict weighing theory, the relevant sense of 'most reason' as it is

discussed here is as a term for ought, not as a statement of substantive commitment to a

weighing-only account of reasons. I shall use 'most reason' and 'ought' interchangeably

from now on in this chapter.[43]

---

[41] Parfit (2001). Pg. 17.

[42] See Ross (1930)

[43] It is important to recognise that some philosophers find the notion of a *pro tanto* or
weighted reason to be basic. As discussed in *The Domain of Reasons* by John Skorupski,
one may not be able to get everything that one needs with only *pro tanto* reasons. This is
particularly evident in the case of reasons for belief. Looking at the different *pro tanto*
reasons, there may be more reason to believe one thing than another. However, the
evidence for both may not be very good, and thus there is not enough reason to believe
either. To remedy this, Skorupski suggests that one needs the notion of sufficient reason
in addition to that of reason. When there is sufficient reason to believe something, then
one ought to believe it, whereas when there is most reason to believe something, there
may not be enough reason such that one ought to believe it.

What we ought to do is, on Parfit's theory, given to us by the facts, regardless of whether or not we are aware of them. The fact that it will be a cold winter gives me most reason to donate my old warm clothes to charity, even if I am ignorant of the long-range weather forecast. The fact that it will save your life is a reason for you to take this pill, regardless of whether or not you know that it will save your life. Or, as in Parfit's example, even though you believe the hotel is on fire, you ought not to jump into the canal, because the hotel is not, in fact, on fire. Particular substantive theories of reasons and oughts will explain how it is that these facts generate reasons and what sort of facts are the right sort of facts.

There is a problem with saying that facts give us reasons but beliefs do not. When I believe that the hotel is on fire, it is a fact that I believe that the hotel is on fire. To be more precise, Parfit needs to say that the fact that we believe something does not *prima facie* give us a reason to act as if the contents of our beliefs are true. Facts about beliefs can still give us reasons. The fact that I believe I am 10ft tall is a reason for me to see a psychologist, but it is not a reason for me to say that I am the tallest man in the world.

In addition to saying that reasons are given to us by the facts, Parfit also denies that reasons are given to us directly by desires. This is an important point in his theory. A desire theorist of reason might agree that facts are, or give us, reasons. However, the desire theorist would in that case hold that the only facts that give us reasons are facts about desires. Furthermore, a desire theorist need not hold that we must know when we have a reason to do something, as there is no requirement that we know what all of our desires are. Although saying that reasons are facts does not formally exclude the

possibility that facts about your desires give you a reason to achieve the contents of those desires, Parfit opposes the view that desires normally give you a reason to achieve what it is that you desire.

In the case of practical normativity, Parfit allows for the possibility that facts about beliefs may be relevant to determining what there is most reason to do. So when he restricts rationality to the domain of beliefs, he means something quite specific. Parfit understands rationality in counterfactual terms: it is most rational to do what there would be most reason for us to do if our non-normative beliefs were true:

> I have rejected the common view that our desires [and actions] are rational when and because they depend on true or rational beliefs. Often, I have said, the opposite is true. Our desires [and actions] are rational when they depend on beliefs whose truth would give us reasons to have these desires. It is irrelevant whether these beliefs are either false or irrational. Remember next that, in making these claims, I have been discussing only non-normative beliefs. When we turn to normative beliefs, we should make different claims.[44]

So, if my belief about the hotel's being on fire were true, then it would be true that I ought to jump into the canal. Therefore, it is rational for me to jump into the canal, even though I ought not to. It is most rational for me to φ just if, were the contents of my relevant beliefs true, it would be the case that I have most reason to φ.

Before moving forward with the discussion of Parfit's work, it is important to look briefly at his notion of 'most rational'. Given his views about the relationship between rationality and reason, the most likely analysis of 'most rational' is that which is

---

[44] Parfit (2001), p. 30.

rationally required. Rational requirements are strict, which means that there are no

rational factors that outweigh them within rationality. Rational requirements may bear

features in common with oughts. For instance, it may be the case that one is rationally

required to φ only if one believes one can φ or does not believe that one cannot φ.[45]

Another possible restriction is that if one is rationally required to φ, then one cannot be

rationally required not to φ at the same time. Parfit is not explicit about these logical

features of rational requirements, nor does he discuss how far the parallel between 'most

reason' and 'most rational' is meant to hold in terms of the logical structures of those

concepts. Parfit may be committed to a kind of weighing structure within rationality,

analogous to a weighing account of reasons, if he believes that reasons themselves have a

weighing structure. Since it is not clear from the text what his views about weighing are

or what logical constraints he wants on most reason, and since Parfit's position on these

matters is not directly relevant to the argument here, 'most rational' will be taken as

being synonymous with 'rationally required'. The logical features of a rational

requirement will be taken to be the same as at least some of the features of *ought*, but

with the important distinction that rational requirement is a strictly rational rather than a

properly normative notion.

2.2 The trouble with rationality

The counterfactual view of rationality leads to some troubling implications. In

---

[45] I am not endorsing or rejecting the view that 'ought' implies 'can', but observing that those who think that it does may also think that rational requirements are subject to a similar constraint.

this section, I shall discuss two problems with the counterfactual view. The first problem is that it may be difficult for Parfit to give an account of practical rationality on this view, when another common assumption is made. There are hints in 'Rationality and Reasons' that Parfit accepts a view that I have earlier called 'normative separatism'[46]. Normative separatism is the view that theoretical and practical reasons are of fundamentally different kinds, and thus, while elements of their logic may be analogous, they constitute two distinct classes of normativity that cannot meaningfully be related or compared. Parfit's account of what there is most reason to do, combined with normative separatism, may prevent him from providing a successful account of what it is most rational to do.

The second problem that this section will address concerns a technical difficulty in Parfit's account. Parfit's account of rationality discusses what we have most reason to do or desire (the reasons for both are identical in Parfit's account). I shall argue that this view of practical reason forces onto Parfit's account the view that we are rationally required to perform certain actions. This is in contrast to an account in which one would be rationally required to intend those actions. An account of rationality in which there are rational requirements to perform actions rather than merely intend them leads to some strongly counterintuitive results.

2.2.1 Rational requirements, beliefs, and intentions

Practical rationality is distinguished from theoretical rationality in that it is concerned with the rationality of action, whereas theoretical rationality is concerned with the rationality of belief. I am assuming that when practical rationality is said to be the

---

[46] See the introduction to this thesis.

rationality of actions, this means that practical rationality also concerns the rationality of some action-related mental state(s). For purposes of the argument here, I am taking intentions to be the mental states that set an agent to perform an action. In general, when one intends to do something, provided that one is not thwarted, one does or at least attempts to do the intended action. Later in this chapter, I shall discuss what would be wrong with a view of practical rationality that is primarily concerned with actions themselves rather than with intentions or other action related mental states.

The core issue addressed in this section is whether or not Parfit's account of practical rationality can deliver intentions. The phrase 'deliver intentions' may be somewhat unclear; I mean the following by it. An account of practical rationality ought to say when it is rational to have or form certain intentions. An account of practical rationality that only says, for example, what one ought to believe about actions or intentions is not really an account of practical rationality. Parfit himself speaks of what we have most reason to do, or desire, or what is most rational to do, or desire. As mentioned in the previous paragraph, I am taking it for granted that intention is the relevant mental state associated with action. In the rest of the chapter, I shall speak at times of what there is most reason to do, or intend, rather than of what there is most reason to do, or desire. The substance of the argument is not affected by the substitution of intention for desire, and making this substitution will bring the discussion in this chapter in line with the discussion throughout the rest of this thesis.

Parfit's definition of what is most rational is that it is most rational to do (or intend) what it would be the case that one ought to do (or intend), were one's non-normative beliefs true. On my gloss, Parfit tells us that given certain beliefs, certain

actions are rational.

As identified in the introduction to this thesis, normative separatism is the view that theoretical and practical normativity are two fundamentally different kinds. There is an equivalent view in rationality, rational separatism. On this view, the concept of rationality required in the domain of theoretical rationality is a different one from that required in the domain of practical rationality, rather than its being the case that the same concept of rationality operates on two different domains. One of the important consequences of this is that you cannot do logic that combines theoretical and practical rational requirements under the same rational requirement operator, as there is no overarching notion of rational requirement that can apply to both the theoretical and the practical, just as with normative separatism you cannot combine the normative claims made by theoretical and practical normativity. To provide a brief example, a rational separatist might accept that it follows from your being rationally required to run and your being rationally required to swim that you are rationally required to run and swim. She might also accept that from your being rationally required to believe that today is Tuesday and from your being rationally required to believe that the Queen is a woman, it follows that you are rationally required to believe that today is Tuesday and that the Queen is a woman. However, what a rational separatist cannot accept is that from your being rationally required to swim and your being rationally required to believe that today is Tuesday it follows that you are rationally required to swim and believe that today is Tuesday. Agglomeration does not hold across theoretical and practical rationality. That is because the rational requirements regarding beliefs and those regarding actions are different in kind, so one cannot agglomerate an practical rational requirement with a

theoretical rational requirement.

In this section, I claim that Parfit's view of rationality, when combined with rational separatism, is unable to account for the rational requirement to intend to do something because of beliefs that we have, at least in one important circumstance. Although there are strong hints that Parfit is a normative and rational separatist,[47] he does not state his position clearly enough in published work to be certain. Thus this section is a conditional attack on Parfit's view of rationality, at least as far as claiming that he cannot deliver intentions from beliefs. Later in the chapter, doubts about his view that do not depend on his being a normative or rational separatist will be raised.

We can begin the analysis of Parfit's view with an example. You wake up in your hotel room one evening and smell smoke. There is only one door leading out of the hotel, and you believe that it may be blocked by the fire. Fortunately, your room overlooks a slow-moving river. Considering the situation, you form the intention to jump into the river, and you do so.

Parfit tells us that jumping in the river is what is rationally required of you if that is what it would be the case that you ought to do, were your relevant non-normative beliefs to be true. If the hotel were on fire, if there were only one exit, and if that exit were blocked by flames, then it would be the case that you ought to jump out the window. So, on Parfit's view, jumping out the window (or intending to do so), is what one is rationally required to do.

To understand why Parfit's account of rationality is difficult to motivate when it is combined with rational separatism, we must examine the logical structure of rational requirements more closely. I shall argue that the rational requirements that could

plausibly be used to justify Parfit's view are not consistent with normative separatism.

John Broome discusses two ways of analysing sentences of the form 'If p, then you ought to q'.[48] The first way is to take the ought as having a *narrow scope*. A narrow scope ought has this logical form ('O' here represents ought): $p \rightarrow Oq$. This ought is a *detaching* ought, as we can detach the conclusion Oq when p is true.

The second analysis of English sentences of the form 'If p, then you ought to q' is as having a *wide scope* ought. A wide scope ought is an ought that governs the entire conditional rather than only the consequent. It takes the logical form: O($p \rightarrow q$). We can say that a wide scope ought is *non-detaching*, as the truth of the antecedent does not make it the case that you ought to do the consequent. O($p \rightarrow q$) is satisfied whenever the truth value of the conditional $p \rightarrow q$ is true and is not satisfied when the truth value of the conditional is false.

Broome's analysis of conditionals governed by oughts can be applied to rational requirements. Consider the following rational requirement. You are rationally required not to believe ~p if you believe p. This can be interpreted as either a wide or a narrow scope rational requirement. The wide scope version would look like this:

WS: Rationally Required[you believe $p \rightarrow$ ~(you believe ~p)]

whereas the narrow scope would be formed like this:

NS: you believe $p \rightarrow$ rationally required ~(you believe ~p)

---

[47] See footnote 43.

Of these two analyses, the wide scope reading is more plausible in this instance. To see

why, consider the fact that there is no information about how rational it would be for you

to believe p or to believe $\sim p$ based on what you believe the evidence to be. In other

words, there is no information about your believed reasons for believing p or believing

$\sim p$. The narrow scope reading tells us that if you believe that p is the case, then you are

rationally required not to believe $\sim p$. But, you might believe that there are good reasons

to believe $\sim p$ and none whatsoever to believe p. Unless the occurrent possession of a

belief is not just a reason to believe[49] that belief, but a strong enough one to make it so

that indefeasibly it is the case that you are rationally required not to believe its negation,

then it is not clear what could motivate the narrow scope reading of the requirement.

By way of contrast, the wide scope reading looks quite plausible. Given that the

requirement as expressed says nothing of whether or not there are believed reasons that

favour p or $\sim p$, all it requires of us is that we do not believe $\sim p$ if we believe p. We can

comply with this requirement either by continuing to believe p and not believing $\sim p$,

believing $\sim p$ and ceasing to believe p, or ceasing to believe p and also not believing $\sim p$.

That seems like the right reading of the requirement, because it is the one that tells us not

to hold a pair of beliefs whose contents entail a contradiction. Detaching conclusions

about the rationality of believing p or believing $\sim p$ may be derived from what we believe

to be reasons for p or $\sim p$.

Keeping the scope distinction in mind, it appears that Parfit tries to argue for a

---

[48] Broome (2000).
[49] Gilbert Harman argues that the fact that one holds a belief does provide a reason in favour of keeping it, although not the very strong kind of reason that would be needed here. See Harman (1999).

narrow scope rational requirement on an action, or intention, when he would be at best entitled to argue that there is a wide scope rational requirement. This scope error occurs as a result of Parfit's view about the relationship of our non-normative beliefs to the actions that we are rationally required to perform. Parfit's view is that it is most rational for you to do what you would have most reason to do if your non-normative beliefs were true. If you believe that your friend is thirsty, then it is most rational for you to offer him a drink. That is because if your friend actually is thirsty, then you have most reason to offer him a drink. In the case in which your friend is thirsty, the conclusion that you have most reason to offer him a drink is a detaching conclusion. Since Parfit gives no hint that 'most rational' is logically different in form to 'most reason', then that it is most rational to offer your friend a drink would detach from your belief that your friend is thirsty.

According to Parfit, the rationality of the non-normative beliefs that we hold does not play a role in determining what we are rationally required to do. Instead, what you are rationally required to do is the action that it would be the case that you ought to do if the contents of those beliefs were true, regardless of whether or not you are rational to hold those beliefs.

In the example given above, it does not matter whether or not you believe that there are reasons for you to believe that the hotel is on fire. It is sufficient that if it were the case that the hotel was on fire, then it would be the case that you ought to jump out of the window of the hotel. Parfit does not provide an explicit justification for his view that you are rationally required to φ when φ-ing would be what there is most reason to do, were your beliefs to be true. And, if one accepts separatism about rational requirements, then Parfit's view cannot be justified.

To understand the difficulty of holding Parfit's view and separatism, consider the

reasons why you might fail to intend to jump through the window when you believe your

hotel room to be on fire. There is a rational requirement that if you believe that you

ought to φ, then you intend to φ.[50] This rational requirement has a wide scope form:

RR1. Rationally required (if you believe that you ought to φ, you then intend to φ).

This requirement provides a link between belief and intention, by specifying the belief

state that requires you to intend to do something. The mental state in the antecedent has a

distinctive content, namely that the content of the belief is of the form 'I ought to φ'. If

you were checking a catalogue of your beliefs for beliefs that would rationally require

you to form an intention, you would know that those of this form would provide such a

rational requirement.

On the other hand, there is nothing distinctive about the belief that the hotel is on

fire. It has the same form as other beliefs about the world. Looking through your list of

beliefs, nothing in particular would stand out about this belief that would tell you directly

that you are rationally required to intend to jump out of the window.

But this belief, at least in combination with some others, may well point you

clearly to the belief that you ought to jump out the window. You believe that the hotel is

on fire. You believe that the normal exit is blocked and that the only other way out is

through the window. You also believe that if you do not leave, you will die, and that

---

[50]   An explanation of why this is a rational requirement is given in 3.2.1. As far as I am
aware, this requirement was first discussed by Jonathan Dancy. See Dancy (1977).
There is also an extensive discussion in Broome (2001*b*).

allowing yourself to die would be a mistake. The fact that the contents of some of your beliefs would make it the case that you ought to jump out a window, were those beliefs true, plausibly gives you a rational requirement that you believe that you ought to jump out of a window. However, there is nothing distinctive about these beliefs that would generate a rational requirement that you intend to jump out of the window.

In order to be led to a rational requirement that you intend to jump, you need a further step. According to the rational requirement that you intend to φ, if you believe that you ought to φ, you would be rationally required (in the wide scope sense) to intend to φ, if you had concluded that you ought to φ.

On this line of reasoning, I could fail to intend to φ, if I believed I ought to φ, owing to an error either of theoretical rationality or of practical rationality. The theoretical error would be to fail to reason correctly such that you had the belief that you ought to jump out the hotel window, when you also had the non-normative beliefs about the hotel's being on fire and so on. This error would be one of theoretical rationality because it involves failing to have a belief that you are rationally required to have in light of other beliefs that you have. You could also fail to form the intention to φ, even though you believe that you ought to φ. This would be an error of practical rationality, as it would involve failing to have an intention that you are rationally required to have.

Parfit's view could be justified by combining the theoretical and practical rational requirements discussed above. Believing that the hotel is on fire (at least in combination with some other beliefs) rationally requires you to believe that you ought to jump out of the window. Believing that you ought to jump out of the window rationally requires you to intend to jump out of the window. So, if we combine the two sets of requirements,

believing that the hotel is on fire (along with some other beliefs) rationally requires you to intend to jump out the window.

A more precise account of what is meant by combining practical and theoretical rational requirements is necessary, as is an explanation of why doing so makes sense out of Parfit's view. As a terminological convenience, I continue to use 'ought' in place of 'most reason' and 'rationally required' in place of 'most rational':

Parfit's view (modified to be about intending) is:

C. (If you believe p, then you are practically rationally required to intend to φ) if and only if (if p, then you ought to φ)

FC. $(Bp \rightarrow PRRI\varphi) \leftrightarrow (p \rightarrow O\varphi)$

In this section, I have argued that a wide scope reading of this rational requirement is correct. The modified conclusion is:

MC. You are practically rationally required (if you believe p, then you intend to φ) if and only if (if p, then you ought to φ)

FMC. $PRR(Bp \rightarrow I\varphi) \leftrightarrow (p \rightarrow O\varphi)$

I have said in this section that there is not a direct rational requirement that from believing that p you are rationally required to intend to φ, when it is true that if p then

you ought to φ. However, there is a theoretical rational requirement when this is the case:

T. You are theoretically rationally required (if you believe p, then you believe that you ought to φ) if and only if (if p, then you ought to φ)

FT. $TRR(Bp \rightarrow BO\varphi) \leftrightarrow (p \rightarrow O\varphi)$

Further, there is a practical rational requirement when you believe that you ought to φ.

P. You are practically rationally required (if you believe that you ought to φ, then you intend to φ)

FP. $PRR((BO\varphi \rightarrow I\varphi)$

This is where the resources of the rational separatist and the rational non-separatist make an important difference. The rational non-separatist does not think that two separate concepts are in use with the practical rational requirement operator and the theoretical rational requirement operator. The rational non-separatist only need have one operator, the rational requirement operator. Because of that, the rational non-separatist can do logic including both practical and theoretical rational requirements. This makes Parfit's view consistent with rational non-separatism. It is not consistent with rational separatism, because rational separatism holds that there are two distinct concepts in use, that of a practical rational requirement and that of a theoretical rational requirement. As

the rational separatist takes TRR and PRR to be two different operators, logic cannot be

done combining *operanda* of a TRR and a PRR. The rational non-separatist can derive

the rational requirement that you are practically rationally required (if you believe *p,* then

you intend to φ) if and only if (if *p*, then you ought to φ) from T and P, as long as the

transitivity of rational requirements is also assumed. The rational separatist cannot.

Consider the argument:

P1. $[T]RR(Bp \rightarrow BO\varphi) \leftrightarrow (p \rightarrow O\varphi)$ T

P2. $[P]RR(BO\varphi \rightarrow I\varphi)$ P

P3. $[RR(a \rightarrow b) \& RR(b \rightarrow c)] \rightarrow RR(a \rightarrow c)$ Transitivity of RR

therefore:

P4. $RR(Bp \rightarrow I\varphi) \leftrightarrow (p \rightarrow O\varphi)$ P1, P2, P3

Rational separatism denies P3, without which P4 is not derivable from P1 and P2.

Rational non-separatism can in principle accept P3. However, rational non-separatism

does not require P3. Thus, while an argument for the transitivity of rational requirements

would be necessary, rational non-separatism does not *prima facie* exclude its possibility,

as rational separatism does.

       The evidence in the literature is not sufficient to tell whether or not Parfit is a

rational separatist. Parfit's view is that reasons for action and reasons for belief are

always different.[51] The difference is that reasons for belief are given by facts when those facts are evidence for the belief's contents, whereas reasons for action are given by facts that make the effects of the action good. While it is certain that Parfit is an evidentialist as far as reasons for belief go, he does not say anything explicit endorsing or denying normative separatism.

Although Parfit's account of what it is most rational to do is cashed out in terms of what there is most reason to do, Parfit is also silent on the logical similarities and dissimilarities between rationality and reasons. Given this, it is not clear whether Parfit's being a normative separatist would commit him to being a rational separatist. What is certainly the case is that Parfit cannot consistently hold his view of what is most rational and also be a rational separatist.

2.3 Practical rationality and normative beliefs

In the previous section, some possible difficulties for Parfit's account of rationality were discussed. Those difficulties arose because of the role that non-normative beliefs play in practical rationality. An additional complication for Parfit's account arises out of his views about the relationship between normative beliefs and practical rationality. Parfit says some suggestive things about the difference between the rationality of non-normative beliefs and the rationality of normative beliefs. In the passages quoted in 2.1, he says that it is the contents of our beliefs### taken as if they are true### that determine what it is rational for us to do or desire, when the beliefs involved are non-normative. However, the role of normative beliefs in practical rationality,

_____

[51] Parfit (2001), pp. 30-32.

according to Parfit, is heavily constrained. Some extensive quotation is required:

> Similarly, imagine if this man believes that agony is in itself worth achieving, or believes that we have no reason to avoid agony of future Tuesdays. These are beliefs whose truth would give him reason to prefer [greater] agony on Tuesday [to less agony on another day]. Similarly, if this man believes that this preference is rational, that is a belief whose truth would make his preference rational. But, even if his preference [for greater agony on Tuesday] depends on these beliefs, that does not make it rational. When our desires depend on such normative beliefs, these desires *are* rational only when, and because, these beliefs are rational. If these normative beliefs are false or irrational, like the belief that agony is worth achieving, or that future Tuesdays do not matter, it is irrelevant whether these are beliefs whose truth would make these desires rational.[52]

and he elaborates further:

> There is, however, another ground for claiming that the rationality of such desires, or of our being motivated by such normative beliefs, depends on the rationality of these beliefs. Our reasons to have these beliefs are very closely related to our reasons to have these desires. In the simple cases, that relation is this. We have some desire because we believe that some fact gives us a reason to have it, and we have this belief because this fact does give us such a reason...
>
> ...Since such desires and [normative] beliefs are so closely related, being in these different   ways responses to the same practical reason or reason giving fact, such desires are rational when   and because the normative beliefs on which they depend are rational.[53]

---

[52] *Ibid.*, p. 30. My comments up to this point have been about actions and not desires, and the coming ones shall be about actions, too. Parfit's claims about desires can, I think, be taken also to be claims about actions. His particular view is, if I understand it correctly, that a reason to desire something is also a reason to act.

[53] Contrast this with the earlier comment 'That may suggest that, in having this desire and this belief, we are responding to the same reason. That is not so. Practical and epistemic reasons are always quite different. But, in this kind of case, the reasons partly

To the put the point another way, there is an overlap here between practical and theoretical rationality. Practical rationality involves, not only responding to our reasons for caring and acting, but also responding to our epistemic reasons for having beliefs about these practical reasons. This other part of practical rationality, which we can call practical reasoning, is a special case of theoretical reasoning: since it is theoretical reasoning about practical principles.[54]

Parfit has now introduced a useful distinction between practical rationality generally and practical reasoning as a special part of practical rationality. The former is broad in its scope, and includes reacting to judgements about the actions we perform or the desires we have. The latter, however, is just a special case of theoretical reasoning### theoretical reasoning about practical principles.

With the above comments and this new distinction, we can formulate a principle for Parfit about when an action is most rational:

V1. An action (or intention) is most rational for an agent if and only if that action would be what he ought to do, were the agent's non-normative beliefs true.

We should be wary of this principle. Parfit indicates later[55] that he wants this principle to allow us to say that an agent is irrational when he acts on irrational normative beliefs. This principle is unnecessary, however, to level the general charge of irrationality at an agent who has acted on an irrationally held normative beliefs. The fact that the belief is

___

overlap…' Parfit (2001), p. 31. Parfit has not carefully spelled out his theory of reasons, and it would behove him to do so, since on some accounts he will have his theoretical and practical reason being one and the same in this case, contradicting his claim that they are always different.

[54] Parfit (2001), pp. 31-32.
[55] *Ibid.*, p. 33.

irrational already makes the agent *theoretically* irrational. It is not clear why her forming

an intention when she has this belief is *practically* irrational. An agent is clearly

practically irrational when she fails to form the intention to φ when she believes that she

ought to φ. As Parfit concedes, her forming the belief that she ought to φ, when that is

an irrational belief to form, is a type of theoretical reasoning error. So, there is no need

for V1 to account for an agent's irrationality when she has an intention deriving from an

erroneously formed normative belief. She is already theoretically irrational, and that

seems like the correct criticism of her.

If Parfit is a normative and rational separatist, then it would be odd for him to

hold that theoretical failures should be able to draw practical criticism, because the

practical and the theoretical are meant to be separate spheres of normativity and

rationality.

There are additional good reasons to reject the idea that an agent is practically

irrational when she acts on a poorly formed normative belief. As cited in the previous

chapter, Harman offers an account of the distinction between theoretical and practical

reasoning. The former, he says, is reasoning, the upshot of which is a belief. And the

latter is reasoning, the upshot of which is an intention (or on Parfit's view, perhaps, a

desire). On Harman's account, nothing need be said about the particular contents of the

beliefs or intentions. Whatever criteria determine correct and incorrect theoretical and

practical reasoning sequences, they will be criteria for correctly moving from beliefs to

other beliefs in the theoretical case and for moving from beliefs (and possibly intentions,

too) to intentions in the practical case.

Taking Harman's view into account, the fact that a piece of theoretical reasoning

is about actions does not make it a species of practical reasoning. The only respect in which theoretical reasoning about actions is practical is the same respect that our theoretical reasoning about food may be said to be culinary. The appellation is purely a matter of the subject of the theoretical reasoning, and so far as I can see, there is no principled reason why theoretical reasoning about action deserves to be called practical reasoning in any stronger sense than reasoning about food deserves to be called culinary reasoning. Or as G.E.M. Anscombe put it, just because the first premise of the reasoning concerns minced pies, it makes little sense to designate it as 'mince-pie reasoning'. It is not practical reasoning in the sense that Parfit seems to intend.

This point has more than a taxonomical significance. 'Most rational' is originally defined by Parfit as being what there would be most reason to do, were our beliefs correct, regardless of whether or not they actually are. Parfit then goes on to specify that this criterion only applies to non-normative beliefs; what is most rational to do cannot depend on irrational normative beliefs. But, if there is no philosophically important sense in which theoretical reasoning about actions, even if it is theoretical reasoning about reasons to act, counts as practical reasoning, then Parfit appears to lack good grounds for applying different constraints to normative and non-normative beliefs in judging the practical rationality of an agent or his actions. Theoretical reasoning is about the relation of beliefs and their grounds. Practical reasoning is about the relation of actions and their grounds. Parfit tells us that the grounds of non-normative beliefs are irrelevant in determining the rationality of the actions that rest upon them, whereas the grounds of the normative beliefs upon which actions depend do matter. If practical reasoning, as Parfit means it, is not genuinely practical### that is if the conclusion of practical reasoning is

just a normative belief### then it seems that some explanation is owed as to why the rationality of a normative belief matters when the rationality of a non-normative belief does not. What is owed is an explanation of why, while theoretical rationality and reasoning are not supposed to determine what it is most rational to do given one sort of beliefs, non-normative ones, the theoretical rationality of another sort of belief, relevant normative beliefs about action, does affect the rationality of an action.

This leaves us very much back where we started. It is not clear what claim is being made when one is judged practically irrational. Failing to do what it is most rational for one to do can indicate one of three types of errors on Parfit's account. Two are errors of theoretical rationality and one is an error of practical rationality. The first error of theoretical rationality is a failure to reason correctly from one's non-normative beliefs to the correct normative belief (that one ought to φ). The second theoretical error is to employ an irrationally held normative belief about a practical principle in reasoning about what one ought to do. In this second case, the error is in holding or acquiring the incorrect normative belief. The practical error is the failure to form the intention to do what one believes one ought to do. Only the latter is a good candidate for being an instance of practical irrationality. Parfit's account of practical rationality, while including that practical case, does not give practical rationality any distinctive and interesting use, because practical rationality is not only practical on his account.

One further point warrants mention. This is the problem of limited resources in cognition. Suppose you hear a noise that sounds like a furnace exploding downstairs in the hotel. You then see what looks like thick smoke seeping through the door and you smell the odour of charred hotel interior. These infelicities briefly distract you from

reading *Naming and Necessity*, but not for long. Aside from forming the belief that you

heard a furnace-explosion-type sound, that you see smoke-appearing stuff coming in

under the door, and that you smell something that smells like charred hotel interior, you

form no new beliefs. Crucially, you fail to form the belief that the hotel is on fire. As a

result, you do not form any intentions or take any actions to save yourself from the fire.

Parfit, I think, would argue, and most of us would agree, that there is something

wrong with you for failing to form the belief that the hotel is on fire. If we accept the

transitivity of theoretical and practical rational requirements, then we may say that your

failing the rational requirement to dive out the window results from your theoretical

reasoning error.

We find an agent who has failed this rational requirement to be irrational because

we feel that the consequence of his beliefs, that the hotel is on fire, was obvious enough

that he is rationally amiss for failing to have deduced it from the evidence at hand.

Parfit's account will successfully pick up such cases. But it is not clear how Parfit's

account will cope with cases that have the same type of failure, but in which the failure

occurs not because the agent is thick, but because the correct relevant conclusion for the

agent to draw from his beliefs requires more intellectual capacity than the average person

has.

Perhaps a certain distinctive type of noise, an *n*-type noise, is produced by all and

only hotel fires. When one hears such a noise, there is a hotel fire. It is not widely

known that an *n*-type noise is produced by hotel fires, but a clever acoustic physicist with

perfect pitch would be able to figure out what the source of an *n*-type noise was without

equipment. Suppose it would take such a person a while, perhaps a full day of thinking,

to realise that an n-type noise is uniquely produced by hotel fires. As it happens, an expert acoustic physicist with perfect pitch finds himself in a hotel and hears an n-type noise. The scientist also realises that it will require some thought to identify its source. Being curious, he begins to reason about what could have produced that noise.

The fact that the scientist hears that an n-type noise is a reason for him to jump out of the window, because the noise is both a necessary and sufficient condition for there being a fire in the hotel. Yet, owing to the complexity of the process of reasoning to the conclusion that this noise is produced by a hotel fire, the physicist is unlikely to realise that the hotel is on fire before it is too late.

Normally we would say in such a case that the scientist was not irrational for failing to jump out the window of the hotel in time. If we thought he was guilty of a failing at all, it would be that he was not bright enough to figure out the implication of the noise more quickly. Nonetheless, on Parfit's view, this scientist would be counted as practically irrational. The fact that the scientist heard the noise gives him most reason to jump out the hotel window. The scientist believes that he heard an n-type noise. When the contents of that belief are true, the scientist has most reason to jump out the window, so his belief that he heard the noise means that it is most rational for him to jump out of the window. His failure to jump out of the window can be attributed to the immense difficulty of correctly identifying the source the noise, slowing his ability to realise its implications for his wellbeing. Intelligence and irrationality are wisely regarded as not being one and the same, yet for Parfit, this scientist will have failed to have done something that he was rationally required to do. This seems too demanding for a theory

of rationality; irrationality and lack of genius should not be confused.[56]

2.3.1 Practical rationality and actions

There is a further worry about Parfit's construal of 'most rational'. By his definition, one acts irrationally when one fails to *do* what it is most rational to do. Throughout this chapter I have primarily been discussing practical rationality in terms of intention. As pointed out at the beginning of the chapter, this is because I take rationality to be a system of relations among mental states. Nonetheless, Parfit himself does not discuss intentions, but rather actions and desires. I do not propose to evaluate his view in terms of desires, as I do not think that desires hold much importance in practical reason, but I do wish to discuss the consequences of taking practical rationality to be about actions rather than intentions. An example will help illustrate the worry.

Believing, correctly, that I ought to do what saves my life if there is no negative effect on others, and believing, incorrectly, that my hotel is on fire and I can save my life by jumping into the canal, I open the window and leap into the canal. I have acted rationally and am not open to rational criticism.

Now imagine that I am in the same position, only when I leap towards the window I am blown backwards by a great gust of wind and never make it out of the room. My action has been thwarted through no fault of my own, and yet I have failed to do what it was most rational for me to do### i.e. jump out the window. So, I have been irrational, and yet not at fault.

This is a worrying consequence of Parfit's taking practical rationality to be about

---

[56] For an excellent explanation of the full force of this point, see Cherniak (1992).

actions. Doing so does not account for ways in which the world can impose itself on our ability to act at various times, and makes us be guilty of rational failures because of consequences outwith our control. Sometimes I can intend to do what is most rational and follow through with an attempt at the action, and yet the world can intervene with the effect that the action does not get performed. In the case of the gust of wind, I did everything correctly, but I had bad luck about the weather.

One possible objection to the gust of wind example is that I could not, at that time, have leapt out the window. As 'ought' is taken to imply 'can', perhaps 'rational', too, implies 'can'; I would only be irrational for failing to act in a given way when it was physically possible for me to do so. Even this restriction is not enough to protect us from saying a person is irrational when she is not. Consider the case where I slip on the carpet on my way out the window, and, instead of leaping out, I bang my head on the window sill and am knocked unconscious. Is there anything irrational about my failure? On the normal use of the word 'rational', there is not. After all, I did not reason poorly, and I did not respond incorrectly to reasons that I had good reason to believe that I had. What I had was a spot of bad luck, but no failure of rationality. In this case, one could not appeal to a rational implies can condition, as I certainly could have jumped out of the window, I just had the bad luck to slip. Nonetheless, it was physically possible for me to jump out the window, despite the fact that I did not succeed at doing it.

Parfit is committed to thinking that actions are proper objects of rational judgement because he thinks that there are reasons to act, and rationality is defined in terms of reasons. Parfit could easily avoid problems that stem from actions being the object of rational evaluation if, for example, he thought that there are no reasons to act,

only reasons to desire to do something or to intend to do something. That option is not open to Parfit, because he believes that our reasons to desire or intend are the same reasons as our reasons to act. My reason to intend to dive out the hotel window is also a reason for me to dive out the hotel window. If I have most reason to intend to jump out the hotel window, then I have most reason to jump out the hotel window. In terms of rationality, this means that if it is most rational for me to intend or desire to jump out the hotel window, then it is most rational for me to jump out the hotel window. At least, this is the case if we take rationality seriously as structurally parallel to reason. But, I can fail to jump out the hotel window because I slip or stumble and not because of any mental mistakes. So, I can be irrational because I slipped or misstepped, and this is not our normal view of practical rationality.

2.4 Conclusion to the critique of Parfit

As long as reasons can apply to actions, counterfactual accounts of the kind that Parfit gives will stake our rationality on whether or not we act successfully. This implies that we may be irrational when we fail to act as is rationally required, even though the failure is through no mental (or even physical) failings of our own. A theory of rationality that only relates mental states will not fall prey to this worry.

The other criticisms set out in this section point out two flaws in Parfit's account of rationality. The first is that on Parfit's view of practical rationality, some actions will be deemed practically irrational that are not in any meaningful sense practical failures, but they are theoretical failures. The second flaw is that Parfit makes a category error in

80

classifying a particular sort of theoretical reasoning, reasoning about what one ought to do, as a type of practical reasoning. Parfit's view is incompatible with rational separatism, and lacks general plausibility.

2.5 Wide scope oughts and rationality

The previous section of this chapter examined Parfit's account of rational requirements or what is most rational. Parfit distinguishes between the notion of what there is most reason to do and what it is most rational to do, and defines the latter in terms of the former. This section examines a different view on the relationship between normativity and rationality, the question of whether rationality is really just a branch of normativity. The most interesting work on this question is in recent papers by John Broome.[57] He argues that at least parts of rationality can be captured using wide scope oughts. In this section, I shall argue that while relations governed by wide scope oughts may be a consequence of certain rational requirements, they cannot themselves express rational requirements.

Broome's claim is that certain instances of rationality or reasoning are correctly captured by wide scope oughts. These were discussed in section 2 of this chapter, but a brief review may be in order. In English we express conditional sentences with normativity attached by the locution, 'If p, then you ought to q'. This locution can express two different logical forms. The first logical form follows the English word order.

F1. $p \rightarrow Oq$

F1 expresses a material conditional with the normativity attached to the consequent. To use Broome's terminology, we can say that the normativity in this case is detaching. We say it is detaching because when the antecedent of this conditional is true, then the consequent is true. Thus, when p is the case, we can detach, by *modus ponens*, the conclusion that it is the case that you ought to q. F1 contains a *narrow scope* ought.

The second logical form contains a *wide scope* ought:

F2. $O(p \rightarrow q)$.

In F2, the normativity does not attach to the consequent, but governs the entire conditional. To use Broome's terminology again, we can say that the normativity here is *non-detaching*. It is non-detaching because from $O(p \rightarrow q)$ and p, one cannot detach the conclusion Oq. In chapter 3, there is a more extensive discussion of the logic of wide-scope oughts. For the purposes of this section, it is adequate to take away two things about wide and narrow scope oughts. The first is that narrow scope oughts are detaching: that one can detach a normative consequent when the antecedent is true, whereas one cannot do so with wide scope oughts. The second is that in the case of wide scope oughts, the agent to whom the ought applies is responsible for the truth of the conditional. Thus, one has satisfied the normative requirement of a wide scope ought as long as any of the truth conditions for the conditional have been met### that the antecedent is true and

---

[57] Broome (1999), (2001*a*), (2001*b*), and (2002).

the consequent is true, that the antecedent is false and the consequent is false, or that

consequent is false and the antecedent is true.

One reason for taking the requirements of rationality to be normative claims, i.e.

to be a disjunction of things for which one has reasons or which one ought to do, believe,

and so on, is that doing so explains the pressure to be rational. If rational requirements

are something other than oughts or reasons, then there is a separate question of why, or if,

one ought to be rational. If the requirements of rationality are just expressions of

particular normative claims, then there is no question as to whether one ought or has

reason to be rational; to ask the further question would be to ask if one ought to do what

one ought to do.

The use of wide scope oughts to explain at least some rational requirements is

important to any effort to give an account of rational requirements in terms of normativity

because of what Michael Bratman calls the 'bootstrapping problem'.[58] The bootstrapping

problem occurs when we try to explain the role of beliefs and intentions in rationality.

Consider my belief that it is Tuesday and also my belief that if it is Tuesday, then

I am in Belgium. These two beliefs rationally require me to believe that I am in Belgium.

If we are giving an account of rational requirement in terms of normativity, there are

three plausible analyses of this particular example. The first analysis is that there is a

narrow scope ought attached to the consequent:

F3. $[\mathrm{B}t\ \&\ \mathrm{B}(t \rightarrow b)] \rightarrow \mathrm{OB}b$

---

[58] Bratman (1999). See especially pp. 24-27.

The second analysis is similar to the first, but with a reason rather than an ought attached to the consequent:

F4. [Bt & B($t \rightarrow b$)] $\rightarrow$ RBb

F3 and F4 are both subject to the bootstrapping objection. I may have no reason whatsoever to believe that it is Tuesday or that if it is Tuesday then I am in Belgium. If both my belief about the conditional and my belief about its antecedent are unjustified or there is no reason to believe them, then it is hard to see what reason there is to believe the conclusion. Just the having of some beliefs does not in normal circumstances stand as evidence for the truth of the logical consequences of those beliefs.[59]

There is a parallel case in practical rationality. Consider what I shall call the *instrumental principle*. This principle tells us that intending an end and believing something to be the necessary means to that end rationally requires that one intend what one believes to be the necessary means. A silly man intends to drop an anvil on his foot. In order to do so, he believes that he must purchase an anvil from the local ironmonger. According to the instrumental principle, this silly man will be rationally required to intend to purchase the anvil. However, this individual has no reason to drop an anvil on his foot. It will not make him happy, satisfy a desire, or benefit him or anyone else. Just having the intention does not appear to bootstrap a reason into existence. Without any reason to drop the anvil on his own foot, it is not clear why there would be any reason for the man to purchase the anvil. So, to say that intending p and believing that p only if q

---

[59] I am assuming that there are no non-evidential reasons for the beliefs in question.

(where q is an action the agent can perform) gives an agent a reason to q looks implausible in cases where there is no reason for p.

The thrust of the objection to using detaching normativity when characterising rationality is that no matter how little reason there is for an agent to have certain beliefs or intentions, or indeed no matter how much reason there is for an agent not to have those beliefs or intentions, having those beliefs or intentions will make it the case that there is some reason for the agent to have the mental states that follow rationally from the initial ones. In other words, if rationality is expressed by detaching normative relations, there would be reason for an agent to believe what follows immediately from two irrationally held beliefs, and there would be a reason for an agent to intend what follows immediately from an irrational intention and belief.

As a result of the difficulties involved in expressing rational relations with detaching normative ones, there is some appeal to the idea that rationality is really a system of non-detaching normative relations, or wide scope oughts. I shall argue that rationality is not such a system.

The claim that rationality is not a system of wide scope oughts needs some clarification, because there are two distinct issues involved. The first issue is whether one can accurately characterise all of rationality as a system of wide scope oughts, or if there is more to rationality than wide scope oughts. The second issue is whether any part of rationality can be accurately captured by wide scope oughts.

To begin with, it will be useful to consider more carefully what notion of rationality we are trying to capture. There are two distinct elements to rationality, only one of which we need be concerned with. The first element of rationality, the one that is

not (or is not very) relevant here, is the descriptive part. Descriptive theories of rationality aim at describing how we reason or try to give an accurate picture of the relations among people's mental states. The second element is the normative part, and it is this one in which we are interested.

Wide scope oughts, called 'normative requirements' by Broome, are normative relations, and the notion of rationality that one might attempt to capture with them is the normative notion of rationality, i.e. what sorts of relations *ought* to hold among an agent's mental states. Again, this is in contrast to what relations *do* hold. To progress this far is just to see that there is an additional thing to specify, and that is which set of relations rationality concerns.

Parfit's view does attempt to specify which relations are the ones that count. These are relations among one's non-normative beliefs, one's desires or actions, and rationally held normative beliefs or perhaps just true normative principles. But, as discussed above, Parfit's view is problematic in a number of ways. In searching for new ways to specify which relations are part of rationality, we might want to see which normative requirements, governing relations among our mental states, are true. When we find a true one (on whatever grounds we make the determination that it is true), we can take it that we have found a rational relation among some mental states.

This method will not work, however, because there may be deviant paths to generating normative requirements among mental states. Sometimes it is good for you that certain relations among your mental states hold, irrespective of their rationality. You might get a big prize for conforming to the conditional that if you believe it is Tuesday, then you intend to go to Belgium, so you ought (if you believe it is Tuesday then intend

to go to Belgium). That normative requirement does not arise from a requirement of rationality, despite its holding between two mental states. The explanation of why the 'ought' holds plays some part in assessing whether the normative requirement is an instance of rationality or is not an instance of rationality. The explanation in this example is that the normative requirement holds because it is good for you that if you have one mental state then you have the other. This is the wrong sort of explanation of why the normative requirement holds for it to be a normative requirement expressing a principle of rationality.

It is probably futile to try to come up with a precise specification for distinguishing between normative requirements among mental states that capture some element of rationality and those that do not. Whatever the correct account is, it will have something to do with the nature of the relations among the contents of the various mental states involved. One might say that these requirements will hold in virtue of certain logical and conceptual relationships among contents of the mental states. These relations can be quite disparate in form, which makes providing a clear set of criteria beyond the scope of this chapter.

The difficulty of providing a precise analysis suggests a point of method, one that I shall follow here. The test for whether normative requirements are adequate in one way or another for capturing our notions of rationality should be to see which, if any, uncontroversial, or at least plausible, instances of rationality are accurately captured by normative requirements. This involves assuming that at least some relations among mental states actually are rational relations. Although this approach would beg the

question when arguing against some views,[60] it should not beg the question here, as it is precisely these sorts of cases that normative requirements would be expected to capture.

2.5.1 Why rationality is not a system of normative requirements

State given reasons, also called attitude given reasons, are reasons for having an attitude or being in a state that are in some respect 'unusual' reasons, or reasons not focused in the typical way on the normal determinants of reasons for that attitude or state.[61] The distinction between state and object given reasons that people have in mind is this. Normally we admire someone because she is admirable### perhaps she has through hard work and great personal integrity achieved a position of great importance. Perhaps she has endured some suffering with nobility and dignity and still finds time to selflessly help others. These are the sort of characteristics that would, under normal circumstances, provide or be reasons for us to admire someone. We can, following Parfit, call these reasons 'object given reasons'. Parfit calls them object given reasons because the reason for admiring something is given by certain properties of the thing being admired. By way of contrast, there might be a very different sort of reason to admire someone. A friendly billionaire might offer you half of his fortune if you will admire Alex. Alex is a lazy sadist with a poor sense of humour. He lacks all the qualities that we might normally consider admirable. Still, the prospect of a large reward for so doing is a reason for you to admire him. This sort of reason is a state given reason, a reason that depends not on

---

[60] For example, one could not start from this position when arguing against the view put forward by Stich (1993), where he claims that our picture of which relations among mental states are rational and irrational is completely mistaken.

[61] For the difficulties involved in giving an accurate account of the distinction between

how well the object of one's attitude 'fits' the attitude, but one that depends on incentives for holding that attitude.

Just as there can be state given reasons for certain affective attitudes, like admiring, and for propositional attitudes like intending and believing , there can also be 'state given' reasons for reasoning in certain ways[62] or for having or not having certain combinations of mental states. An example will make this point clearer. This example makes use of a paradigm principle of rationality, that believing a conditional and its antecedent rationally requires you to believe its consequent. This may be expressed as the following normative requirement:

NR1. $O\{[Bp \ \& \ B(p \rightarrow q)] \rightarrow Bq\}$

Note that the ought governs the entire conditional rather than the consequent, so we cannot detach OBq on account of the antecedent's being true.

Some other eccentric billionaire might offer you a prize for believing p, believing if p then q and yet not believing q. According to our normative requirement, NR1, above, you would not be as you ought to be should you have the first two beliefs and not have the third. NR1 is supposed to be the claim that it is normatively required that the following conditional be true of you: if you believe p and believe if p then q, you believe q. Imagine now that with the money you will get from believing p, believing if p then q

state and object given reasons, see Rabinowicz and Rønnow-Rasmussen (2003).

[62] In the case of reasoning, 'state given' is a terminological infelicity. Reasoning is a process, not a state. Perhaps with processes, this class of reasons ought to designated 'process given'. However, to avoid introducing additional terminology, I shall continue to use 'state given' when referring to such reasons for processes.

and not believing *q*, you could and would feed all the hungry people in the world. We might very well think that you ought, assuming you can, to believe *p*, if *p* then *q*, and not believe *q*. You are, in this case, as you ought to be. We can express this as a new normative requirement:

NR2. O~{[BP & B(P → Q)] → BQ}

On a fairly standard account of the semantics of ought, *ought not* implies *not ought*. This means that NR1 and NR2 jointly entail a contradiction, so one or the other must be discarded.

Intuitions about which one of the two to discard might go either way, but it strikes me, at least, that it would be quite hard to explain how it is that saving all the starving people in the world does not have deontic priority over violating a principle of rationality. If this is the case, then at least sometimes NR1 is false. That we sometimes ought not to have our collection of beliefs be such that when we believe *p* and believe if *p* then *q*, then we believe *q* has a great deal of significance for how we interpret the claim that we are rationally required to be in a state where we do, when we believe *p* and believe if *p* then *q*, then believe *q*.

One way of interpreting this situation is that sometimes we deem rational certain relations among our beliefs, which we normally deem irrational. This interpretation carries with it a rather substantive view that there is more to rationality than the normal constraints of logic on what it is rational for us to believe or intend. I mean to use 'logic' here in a very weak sense, one that includes the sorts of considerations underpinning the

instrumental principle. We ought not to be happy with this interpretation of the situation for two reasons. The first is that we risk losing the notion of rationality when we allow conditions unconstrained by the logical (in a weak sense) relations among the contents of our mental states to be set on which relations among mental states we deem to be rational. The second is that this interpretation seems to conflate two distinct notions### what is rationally required or permitted and what is normatively required or permitted. To use Parfit's language, we could describe the error as the conflation of what is rational with what there is reason for. Of course, the two may overlap or even be coextensive, but it seems to be a mistake to think that 'ought' and 'rationally required' are just interchangeable.

So we can try another way of interpreting the situation, and I think that this one is the most sensible. On this view, the initial error was in thinking that normative requirements correctly expressed principles of rationality. The truth of any particular normative requirement is, like any other ought, determined by the various features of the world on which normativity is dependent. Eccentric billionaire examples can be generated for any requirement of rationality that could be expressed as a normative requirement. Sometimes it is the case that it would be so bad to be rational that one ought not to be.

Rationality, on the other hand, is dependent on some complex of the logical features of the relations among an agent's mental states and their contents. That it would be extremely bad for an agent to have mental states that would allow the agent to be deemed rational does not affect the truth of the matter about whether or not the agent's mental states are such that he may be deemed rational. How an agent *ought* to be

mentally can be a separate issue from whether or not that agent is mentally as rationality requires or permits.

To see how what is rationally required and what one ought to do can come apart, consider prudence and morality. There are many different sorts of practical reasons. Some of these reasons are moral, some are prudential, some may be aesthetic. The explanation of what I ought to do, all things considered, certainly can make reference to reasons of many different sorts. Perhaps I ought not to purchase the painting and instead ought to give the money to charity. This is so because, although there were strong prudential reasons to purchase it, given the pleasure that I would derive from gazing at it, those reasons were outweighed by a moral reason to donate the money to charity. What one ought to do is at least in part determined by how good the outcomes of the possible actions are. This is most definitely not the case with practical rationality. The most rational thing for one to do may very well not produce the best results by any light. This is because we generally take rationality to be a sort of internal matter, something that is determined by the relations of our mental states to each other, and depending on what our mental states happen to be, we cannot necessarily expect that their being well-ordered will result in things going well for us.

All that has been shown so far is that rationality and normativity can come apart, but that is sufficient to cast doubt on the identity of rational requirements and normative requirements. Normative requirements may very well correctly express the normativity of rationality when the circumstances are right for rationality to be normative, but it is a confusion to think that the requirements of rationality are to be *identified* with normative requirements. What normative requirements might, and indeed probably do, capture is

92

the logical structure of some important part of the normativity of rationality when rationality is normative. But even here we must be careful, as there will be various principles of rationality, the normativity of which (when it applies) cannot be expressed with normative requirements. One such case is the requirement of rationality that one ought not to believe a contradiction. When normative, the rule that we should not believe a contradiction is properly stated as: you ought not to believe (*a* & ~*a*). This is not a wide scope ought in form. So, at best we can say that normative requirements do accurately express the normativity of some rational principles, when such principles have normative force.

Conclusion

The aim of this chapter has been to explore the relationship between normativity and rationality in two distinct ways. The first section of this chapter looked at Parfit's account of rationality, in which *most rational* was cashed out in terms of *most reason*. I argued that a Parfit-style account requires some justification and that justification for such an account will rely on a denial of rational separatism. I also argued that Parfit's account of rationality treats theoretical rationality when non-normative beliefs are involved differently from theoretical rationality when normative beliefs are involved, and that this difference in treatment is unsupported. Rational separatism, combined with the problems in Parfit's account of the different status of normative and non-normative theoretical reasoning, undermine his ability to give an account of rationality using counterfactual claims about reasons.

In the second part of the chapter, I argued that rationality must be distinguished from normativity. Looking at Broome's characterisation of some principles of rationality as normative requirements, I argued that there will be counterexamples such that taking rationality to be, at least partially, a system of normative requirements will lead to normative dilemmas.

Appendix to Chapter 2: A problem with counterfactual theories of rationality

A detailed treatment has been given here of possible problems that can arise for Parfit's theory of practical rationality, especially when paired with rational separatism. There is a general worry for counterfactual theories of rationality that define 'most rational' as that which there would be most reason to do, were one's non-normative beliefs to be true. An example brings out the problem.

An eccentric billionaire will give you a large prize if the following two conditions are met. The first condition is that your belief about the colour of his dog's eyes be true. The second condition is that you act inappropriately at dinner tonight. The trick with this particular opportunity is that the billionaire has not told you about it, and you neither believe that you should behave inappropriately at dinner tonight, nor do you believe that there is any reason to believe that you should behave inappropriately at dinner tonight.

According to Parfit's view, regardless of what colour you think the billionaire's dog's eyes are, it is most rational for you to behave inappropriately at dinner tonight. If your belief about the colour of the dog's eyes were true, then you certainly would have most reason to behave inappropriately at dinner. After all, if your belief about the colour

of the dog's eyes were true, behaving inappropriately at dinner would win you a large prize. Yet, it seems implausible to think that it is most rational for you to behave inappropriately at dinner, especially when you have no awareness that you might benefit from doing so.

To put this difficulty in more general terms, counterfactual theories will be vulnerable to situations in which the fact that your beliefs are true itself is an important factor in determining what you ought to do. The counterfactual theorist will have to find a way to limit which facts in the world would count in the counterfactual scenario to avoid cases like these, in which an agent can do something that seems totally without basis, even to the agent, and yet be doing what is rationally required of him.

The problems with counterfactuals do not end there. Two more bear mentioning.[63] The first is a standard problem with counterfactuals, that in the nearest possible world in which the counterfactual is true, other things may have changed. Consider the case in which you believe that the hotel is on fire, but it is not. According to Parfit, the most rational thing to do is to jump into the canal outside your window, because that is what there would be most reason to do, were your belief that the hotel is on fire to be true. It might be the case that in the nearest possible world in which your belief that the hotel was on fire is true, the canal is dry, or some other event occurs making it be the case that there is not most reason to jump out of the hotel window.

The second problem for Parfit's theory is that it does not account for cases in which two of your relevant beliefs are contradictory. Imagine a lottery player who believes that the odds of his winning the lottery are 1 in 100. That same lottery player also believes that he is more likely than not to win the lottery. If the first belief were true,

then the gambler would have most reason not to play the lottery. If the second belief were true, then the gambler would have most reason to play the lottery. This puts Parfit's theory in the awkward position of making it most rational to play the lottery and most rational not to play the lottery.

However, there is a worse problem that arises from having contradictory relevant beliefs. Using the example above, we can set up the relevant counterfactual claim. If your odds of winning the lottery are 1 in 100 and you are more likely than not to win the lottery, then there is most reason to φ. In this case, if the counterfactual is true, then there is a contradiction. From a contradiction, anything follows. So, there would most reason to do anything if your beliefs were true, and therefore anything you do would be the most rational thing to do.

---

[63] John Broome first brought these to my attention.

Chapter III: Normative Conflicts

Introduction

This chapter discusses putative normative conflicts. These putative conflicts may be real, or they may be apparent, and they arise in a variety of circumstances. By 'conflict' here, I mean a situation in which not all of the requirements of normativity can be satisfied, whether it would be logically, physically, or psychologically impossible to comply with two or more normative requirements.   Drawing on the material in chapters 1 and 2, this chapter presents some particularly difficult putative normative conflicts. The purpose of so doing is to show what sort of problems normative theorists must address. Whether the best approach is to accept that there are normative conflicts, or whether it is to take the conflicts only to be apparent conflicts and to find resolutions, this chapter will stress that there are putative normative conflicts. Chapter 4 will discuss some approaches to dealing with the putative normative conflicts laid out in this chapter.

Before setting out the conflicts, some technical terms must be introduced. Two terms that I shall use from time to time in the rest of the thesis are *inter-normative* and *intra-normative*. Inter-normative conflicts are those that occur between two or among more than two domains of normativity. A domain of normativity is defined by the nature of the verb or proposition that the normativity governs### whether it is an action verb, belief verb, feeling verb, and so on, or a proposition that concerns an action, belief, and so on### rather than the source of the normativity (goodness, truth, etc.). For example, theoretical normativity is the domain of normativity that is concerned with beliefs or

belief propositions, and practical normativity is the domain of normativity that is concerned with actions and related mental states. Sometimes, I shall argue, we cannot both do what we ought to do and believe what we ought to believe. This would be an inter-normative conflict, as a requirement of practical normativity conflicts with a requirement of theoretical normativity.

Intra-normative conflicts are those that occur within a particular domain of normativity. Thus, following from chapter 1, if we admit the possibility of non-evidential normative reasons for belief, then we may find all and only non-evidential reasons weigh up towards believing *b*, while all and only evidential reasons weigh up towards believing not *b*.

A technical note on the distinction between intra-normative and inter-normative conflicts as it has been presented needs to be added. Normative operators may range over propositions that involve both action and belief. For example, perhaps some fact makes it the case that you ought to perform action *a* and have belief *b*, where the ought governs the conjunction of the propositions: you perform action *a* and you believe *b*. If there are oughts that govern conjunctive propositions of this form, those oughts are part of a domain of normativity that includes both types of propositions.

This chapter sets out a variety of examples of intra-normative and inter-normative conflicts to show just how broadly normativity is infected with conflicts of a difficult nature. While in chapter 4 I argue that there is reason to take all putative normative conflicts as only apparent normative conflicts, in this chapter the realness or apparentness of normative conflicts is not discussed. The aim here is to show what sorts of difficult conflicts normative theorists must either be prepared to accept or to resolve.

3.1 A note on rationality and normativity

Chapter 2 examined the distinction between rationality and normativity, and how the former cannot be cashed out strictly in terms of the latter. Normativity is concerned with what sorts of things there is reason for you to do, believe, feel, and so on, and what sort of things you ought to do, believe, feel, and so on. Rationality has to do with the correct relations among mental states, as determined by the relations among their contents and the features of the states themselves. In chapter 2, I argued that these relations cannot be given strictly in normative terms, but that we must instead have some independent notion of rationality. This is worth noting insofar as the sense of 'correct' employed in describing rationality is a notion of *rational* correctness, not of normative correctness. Some rational relations depend on the underlying logical relations among the contents of the states. For example, believing *b* and also believing not *b* may be considered irrational. This is because the contents of these two beliefs entail a contradiction, *b* & ~*b*. One is rationally required not to have beliefs that directly entail a contradiction. Other sorts of rational requirements, although dependent on the relations among the contents of one's mental states, may not be so easy to explain. It seems a very basic principle of practical rationality that one is required to intend to φ when one believes that one ought to φ. If there is a strictly logical explanation of this requirement, it is rather less straightforward than the one that explains why we are rationally required not to hold beliefs that directly entail a contradiction. Nevertheless, both rational requirements at least are a function of some relation holding among the contents of an

agent's mental states.

In the previous chapter I discussed two accounts of rationality that defined or explained rational requirements in terms of normativity. While that chapter suggested that an account of rationality cannot be given strictly in terms of normativity, it is still plausible to think that there is a normative requirement to comply with your rational requirements.

In this chapter I shall argue that there is a putative normative conflict that can arise if there is a normative requirement to comply with your rational requirements. The view that there is a normative requirement to comply with your rational requirements differs from the view that rational requirements are normative requirements. On the former view, the concept of rational requirement is not a normative concept, whereas on the latter view it is.

The discussion of the normativity of rationality is spread across chapters 2,3, and 4. In section 2.5, arguments are produced to show that rational requirements are not just a subset of normative requirements. In chapter 4, a more in depth discussion of the normative status of rational requirements is provided. In this chapter, the initial normative status of rational requirements is discussed, as are the putative normative conflicts that arise if rational requirements entail normative requirements.

Organising the discussion of rationality in this way is not as tidy as might be ideal, but I believe it is the most effective way of presenting the material. There is some difficulty involved in motivating the fuller discussion of the normative status of rationality without first examining one of the implications of accepting that rational requirements entail normative requirements, that implication being that putative

normative conflicts arise. So, in addition to its primary function of examining putative normative conflicts, this chapter in part serves to motivate the later discussion of the normative status of rational requirements. Limited discussion of the normative status of rationality must do at present.

For now, it will be assumed that you ought to do, believe, feel, and so on, what is rationally required of you. This thought, that a rational requirement generates an ought rather than a reason rests in the structure of rational requirements themselves. Before moving on to discuss why rational requirements generate an ought rather than a reason, it will be helpful to look at the structure of rational requirements.

One thought is that rational requirements are requirements rather than leanings or some sort of weighing entities; one need not take other things into account to know whether or not one is irrational when one fails a requirement of rationality. If something is rationally required of you and you fail to do it, you are, at least in that instance, irrational; no additional factors enter into consideration in a weighing role. You are rationally required not to believe contradictions; if you believe a contradiction, then you are irrational in some important sense.[64] Of course, rationality may involve weighing. For example, a consequentialist may argue that we are rationally required to do whatever we believe is best. In order to determine what is best, there may be some rational process for weighing up evaluative 'evidence' and drawing a rational conclusion about what is best. In at least some sense, we can say that this would count as using weighing in

---

[64] Although this position is relatively uncontroversial, the level of agreement decreases rapidly concerning broader requirements on the consistency of one's beliefs. Harman believes that there are times when it is rational to entertain two beliefs, the conjunction of which is a contradiction. See Harman (1999). Others hold that weaker inconsistencies may be acceptable. For the latter position see Nozick (1993), Pollack (1991) and Raz (2000). See also Cherniak (1992) for an in-depth discussion of feasible limits on the

rationality.

However, rational requirements, such as the one against believing contradictions, do not appear to arise as a matter of weighing. Believing a contradiction is rationally forbidden (by 'forbidden' I mean 'rationally required not') because we take contradictions to be false, and an agent takes, in some sense, the contents of her beliefs to be true.[65] That the contents of a belief are a contradiction makes it logically impossible for that belief to be true. That the contents of a belief are a contradiction is not one epistemic factor to weigh up against things like empirical evidence; it signals outright that the belief must be false.

One might call this sort of requirement 'strict' in the sense that it is not the kind of thing that gets weighed up against other considerations. Rational requirements in general have this character. Consider beliefs about necessary conditions. If one believes that *p* is a necessary condition for *q*, and one also believes *q*, then failing to believe *p* is rationally forbidden. Alternatively, believing *p* is rationally required when one holds the beliefs that *p* is a necessary condition for *q* and that *q* is the case. This rule derives from the truth conditions for *modus ponens*. From [(*q* then *p),* and *q*] deduce *p*. Although all sorts of factors may go into deciding whether or not any of the premises of a *modus ponens* argument are true, that some conclusion follows by *modus ponens* from two premises is not something to be weighed against other considerations. In cases of reasoning about necessary conditions, the logical strictness of *modus ponens* is reflected by the strictness of the rational requirement. The relation among the contents is not merely a

---

requirements of rationality for human agents.

[65] This may be an indelicate phrasing of the relationship of belief to truth. For an interesting discussion of this relationship, see Velleman (2000), ch. 11, and Wedgwood (2002*b*).

consideration for believing *p* when you believe if *q* then *p*, and *q*; it also puts a rational requirement on you to believe *p* or to cease believing one or both of the *p*-entailing beliefs.

In the case of normativity, this sort of strictness is the character of ought rather than of a reason. Failing to do what there is a reason to do does not, by itself, suggest any kind of normative failure. On Tuesdays in the morning, a lorry comes to collect the rubbish. That is a reason for me to put the rubbish out on Monday evenings, as that way I shall not forget and there will be no need to wake up early to do the chore. On the other hand, there are some raccoons who live in a green area not far away, and leaving the rubbish out overnight is likely to attract them. Raccoons are a great nuisance, and the increased likelihood of them spending time near my house is a reason not to put the rubbish out at night. Weighing up these considerations, there is more reason for me not to put out the rubbish at night, even though there is certainly a reason for me to put it out. Not putting the rubbish out does not constitute a failure on my part to comply with normativity, even though there was some reason for me to put out the rubbish. Reasons are what one might call 'slack'. They may act to commend a certain course, but more normative information is needed before we know if complying with a reason constitutes a normative success or a normative failure (as would be the case if I left the rubbish out overnight).

The question now arises as to what sort of normative force rational requirements have, if they have any at all. I shall claim that if rational requirements have normative force, then they at least give us oughts, although they may give us reasons, as well. This claim requires support, which will be given in the next section of this chapter. Note that

it is a further claim that rational requirements do actually have normative force.[66]

Although it is not the purpose of this chapter to argue that rational requirements do have

normative force, two very general considerations are still worth mentioning. The first is

that if one thinks that rational requirements in general have no normative status, then it is

not clear in what sense there is anything wrong with irrationality. Typically it is thought

that someone who is irrational, or at least persistently irrational, is not as he ought to be.

Part of that feeling may come from the practical difficulties involved in coping with

persistently irrational individuals, but that would not go all the way towards explaining

why people who have serious deficits of rationality are typically considered mentally ill

and it is judged that something is wrong with them### i.e. why they are deemed not to be

as they ought. When we are told that a colleague is highly irrational, this information is

not offered only as a warning, but also as a kind of criticism. We may be wary of the

irrational colleague, but we are also likely to believe that 'something is wrong with him',

i.e. he is not as he ought to be.

The second very general consideration in favour of rationality's having some

normative force is that quite apart from rational considerations, truth seems to be an

important source of normativity. The truth is seen as something worth getting at,

preserving, promoting, and fleshing out. In both practical and theoretical normativity,

truth appears to play an important role. On the practical side, that someone has a false

belief is a reason to inform him of his error, even if nothing of importance hinges on the

---

[66] Stich (1993) provides a powerful attack on the notion that rationality as it is discussed here has any intrinsic normative import at all. Barry Loewer replies from a pragmatist perspective (see Loewer (1994)) with a defence of the normativity of rationality (or to be precise, of truth-directedness in our theoretical reasoning). Pragmatist attacks and defences on the normativity of rationality will be considered in chapter 4.

belief. In traditional epistemology, that something is true or that there are indications that it is true is a reason to believe it. If truth is indeed normative in some way, then it would be surprising for rationality, or at least parts of it, not to have some normative status. There are a large number of rational requirements that operate on the logical relations among the contents of one's mental states. Rationality plays an important role in truth preservation (when one's beliefs are, in fact, true), allowing for the possibility of non-serendipitous truth preservation in one's reasoning. Rationality provides a system whereby truth *could* be preserved in our reasoning, allowing one to move from one group of true beliefs to some other true beliefs. One way of diagnosing what is wrong with persistently theoretically irrational people is that even if they have true beliefs about the world to begin with, they lack the capacity to reason from these beliefs to correct new ones, except by chance. Rational requirements are often connected very closely to truth, and if truth is an important source of normativity, then it is not implausible to think that rationality, too, has normative force.

The importance of truth in rationality is not limited to theoretical rationality. There is a rational requirement that you not both intend to φ and believe that you will not φ. This requirement derives from the relation of intention to truth. When you intend to do something, you set yourself to make it true.[67] There is something inconsistent about setting yourself to make something the case and believing that it will not be the case.

Accepting these considerations, it appears more plausible that there is some normativity to rationality rather than none. The first consideration lends some support to

---

[67] I use this expression advisedly; clearly one need not have the concept of truth to have intentions. I mean only to draw a parallel with belief and the way in which one may be said to take one's beliefs to be true.

the view that rational requirements generate *prima facie* oughts, although that support is not decisive. When one has a *pro tanto* or weighing reason for something, it is only the case that one ought to do it if there are no other reasons that outweigh the first. There is a reason for me to eat ice cream; the reason is that doing so will make me happy. Yet, it is not a default assumption that there is something wrong with me whenever I do not eat ice cream. Facts about my health, my pocketbook, and my need to devote attention to other things are often reasons that outweigh the reason for me to eat ice cream. Rationality is not like eating ice cream in this respect. The default position is that I am not as I ought to be, when I am (persistently) irrational. In general, it seems that you ought to be rational, and it is only in unusual circumstances that you ought not to be rational. Of course, this default position might be taken as suggesting that rather than its being the case that you *prima facie* ought to comply with your rational requirements, there is instead a very strong reason to comply with your rational requirements, a reason that is not easily outweighed. In 4.3.3 and following, a weighing interpretation of the normativity of rational requirements is discussed. For the purposes of this chapter, only the consequences of taking rational requirements to generate *prima facie* oughts will be considered.

3.2. Armed robbery

Derek Parfit used the example of 'Schelling's Answer to Armed Robbery' in his book, *Reasons and Persons*. Parfit used the case as an example of 'rational irrationality', by which he meant that there are times when it is rational to *cause* yourself to be

irrational.

Here a modified version will be introduced for purposes slightly different to Parfit's. In the current example, the intention is to show ways in which taking the requirements of rationality to provide oughts sets up a putative normative conflict. The fallout from this conflict will be discussed in greater detail in chapter 4, but for the moment I shall just make the case for the existence of the putative conflict itself.

The example here is intended to show that there are circumstances in which one ought not to be rational in the very particular sense that one ought not to do or to be as is required by rationality. Yet, requirements of rationality, we are supposing, provide us with oughts### we ought to do or to be as rationality requires. If there are some circumstances in which normativity deriving from moral or prudential considerations requires us not to act (i.e. we ought not to act) in a manner that is consistent with the requirements of rationality, then there appears to be a normative conflict between what morality or prudence might require and what rationality requires . In other words, it seems that normativity may require us both to comply and not to comply with a rational requirement. Below is a modified version of 'Schelling's Answer to Armed Robbery', which is an example of just such a situation.

You see a burglar break into your house, so you phone the police. Because you live in the country, it will take them 20 minutes to arrive. The burglar knows this, and has chosen your house because of his good prospects for a getaway. The burglar is masked, but he has parked his car in front of the window, and you have seen the number plates on his car.

The burglar asks you for the combination to your safe so he can steal the gold in

it. If you give him the combination, he will no doubt kill you, because you have seen his number plates and can tell the police. If you fail to tell him, he will kill the members of your family, whom he has locked in the next room, one by one until you tell him. Then he will kill you, even if you do not tell him, because you can identify the number plates.

Telling the burglar the combination means that you will die, not telling the burglar means that you and your family will die. Barring the possibility of physically overcoming the burglar, there is one solution that may be available to you. If you are irrational, at least in one particular way, then the burglar will be unable to gain the combination to your safe, except by luck. The rational requirement involved, and its normative status, require some more detailed discussion.

3.2.1 The normative status of the instrumental principle and a putative conflict

The way in which you must be irrational is that you must violate the requirements of what I shall call the *instrumental principle*. The instrumental principle is a principle of practical rationality, and it requires you to take what you believe to be the necessary means to ends that you intend. So, if you intend to eat a Pot Noodles and believe that removing the lid is a necessary means to eating the noodles, you can comply with the instrumental principle in one of three ways: by intending to remove the lid to the Pot Noodles, by rescinding your initial intention to eat the Pot Noodles, or by changing your belief about the necessary means to eating a Pot Noodles. What constitutes a failure to comply is retaining your initial belief and intention while failing to form a new intention to remove the lid to the Pot Noodles.

The instrumental principle is justified by the logical relations of the contents of the relevant mental states. John Broome provides a thorough account of this elsewhere,[68] and a more in-depth discussion follows below, but we can note at least casually that the instrumental principle is crucial to allowing people to get things right without being lucky. Assuming that we have reasonably accurate beliefs about means and ends in general, intending a particular end without intending what we believe to be the necessary means would not get us very far. In the case of the Pot Noodles, I would be permanently frustrated in my efforts to have dinner, were I not to comply with the instrumental principle, because at no point would I make the connection between intending to eat a Pot Noodles and intending to take the lid off of the container. That I recognised that taking the lid off of the container was a necessary means to eating the Pot Noodles would not lead to my forming the intention to do so. A crucial link would be broken in my ability to carry out my intentions.

The very basicness of the instrumental principle to rational action is what might allow you to escape from the armed robber by failing to reason in accordance with the instrumental principle. If you fail to reason in accordance with the instrumental principle, the burglar is very unlikely to get the combination from you by killing your family. If you intend to save your family and believe that to save them you must give him the combination, you will *not* form the intention to give him the combination, at least not as a result of intending to save your family. This is because believing that giving him the combination will save your family and intending to save your family would not produce the intention to give over the combination. Coercion loses its effect in this circumstance, and a burglar who recognised that you were in this state would do best to

---

[68] Broome (2000).

give up.

Having lost the ability to coerce you, the burglar is now faced with an inability to get your gold. There is nothing he can convince you of that will then get you to form the right intention (without his being lucky, anyway). And, as an additional benefit to failing to be instrumentally rational, if you are really instrumentally irrational, then the burglar might think that you will not report his number plates to the police. After all, you will no doubt intend to have him captured and believe that the necessary means to doing so is reporting to the police his number plates. Your instrumental irrationality will mean that you will not form the intention to give the police the burglar's number plates, at least not as a result of your intending his capture. Thus, the burglar is best to let you alone and flee.[69]

We can look at the instrumental principle in more detail to see formally why, if you fail to follow it, the burglar will not be able to coerce you into giving him the combination to your safe. The instrumental principle requires you to take what you believe to be the necessary means to an action you intend, or to rescind the intention. It is most plausibly formulated with a wide scope ought governing a conditional sentence.[70] As mentioned in 3.1, the claim that this sort of requirement of rationality, if it has any normative force at all, produces an ought, requires some justification. One loose consideration was provided in 3.1. A more thorough justification is provided here.

The instrumental principle tells us that we are rationally required to take what we believe to be the necessary means to ends we intend. Why this is the case becomes

---

[69] In order to avoid the claim that we ought to *act* as if we do not follow the instrumental principle rather than actually not follow it, imagine that the burglar would somehow see through your attempt to act.
[70] See 2.2.1 for a more thorough discussion on the scope of oughts.

apparent when looking at the contents of the mental states involved. Consider the Pot

Noodles case, which we can formalise the Pot Noodle reasoning as follows:

1. I intend that I eat the Pot Noodles.

2. I believe that my removing the lid is a necessary condition of my eating the Pot

Noodles.

If we only look at the contents of the intention and the belief, we get:

3. I eat the Pot Noodles

4. If I eat the Pot Noodles, then I remove the lid of the Pot Noodles.

From 3 and 4 one can derive 5 by *modus ponens*:

5. I remove the lid of the Pot Noodles.

It is because 3 and 4 entail 5 that there is good reason to think that the instrumental

principle is true. 1 states that I intend to eat the Pot Noodles, and 2 states that I believe

that I must remove the lid in order to eat the Pot Noodles. Lacking the intention to

remove the lid puts me in the position of not having the contents of my mental states

related in the way that is dictated by the logic of their contents.

This rational requirement takes a wide scope form. The reasons for taking the

instrumental principle as a wide scope rational requirement parallel some of those given

in the discussion in 2.2.1. As far as the Pot Noodles example goes, there are no reasons given for or against my intention to eat Pot Noodles or my belief that in order to eat Pot Noodles, I must remove the lid. So, at least as far as the example goes, there is nothing wrong with having or not having the intention to eat the Pot Noodles, believing or not believing that in order to eat the Pot Noodles, I must remove the lid, or having or not having the intention to remove the lid. The only thing that is clearly wrong is intending to eat the Pot Noodles, believing that I must remove the lid to do so, and also not intending to remove the lid. Because the requirement concerns the status of the trio of mental states in this way, the scope of this rational requirement is wide:

I am rationally required [(if I intend to eat a Pot Noodles and believe that in order to do so, I must remove the lid) then I intend to remove the lid].

A narrow scope rational requirement is not appropriate in this instance. If the requirement were a narrow scope one, then it would say that when I intend to eat a Pot Noodles and believe that in order to do so, I must remove the lid, then I am rationally required to intend to move the lid. However, it may be irrational to have the initial belief or intention, in which case it is not clear why it would be rationally required of me to form the intention to remove the lid.

   Although it might be tempting to think that settling on the form of the rational requirement allows us to infer that the normative requirement that the rational requirement generates is of the equivalent normative form, the relation between rational requirements and normativity is not so straightforward.

Some of the discussion of the correct normative treatment of rational requirements will occur in chapter 4, as it is difficult to motivate without first having looked at the putative normative conflicts that can arise from the treatment in this chapter. Here I shall give an account of why it is plausible to infer that a wide scope rational requirement implies a *prima facie* wide scope ought. This is not to exclude the possibility that other normative consequences follow from something's being rationally required. In particular it is not to preclude the possibility that a wide scope rational requirement can provide normative reasons. But discussion of these matters must wait until the next chapter.

One important comment is required before proceeding. In the last chapter, I argued that rational requirements are not just normative requirements of some form.[71] That was not to say that there may not be some normative requirement that follows from your having a rational requirement, it is just that the rational requirement and the normative one are not the same thing.

A wide scope ought may be derivable from a wide scope rational requirement because one runs into the same problems with detaching relations on the normative side as one runs into on the rational side. This is best illustrated with an example. Returning to the Pot Noodles case, eating too many pots of Pot Noodles is not healthy. Suppose I have already exceeded my weekly allowance of Pot Noodles. Next to my Pot Noodles on the shelf is a Pot of Health. Pot of Health supplies me with all the nutrients that I missed out on this week by eating only Pot Noodles.

Suppose that this wide scope rational requirement### that if I intend to eat Pot Noodles and believe that a necessary means to doing so is removing the lid, then I intend

to remove the lid### generated a detaching normative requirement. This normative requirement would be of the form: If I intend to eat Pot Noodles and believe that removing the lid is a necessary means to doing so, then I ought to remove the lid.

However, what I ought to do tonight is eat a Pot of Health. Removing the lid to the Pot Noodles plays no role in my eating a Pot of Health. It looks like the detaching ought will not be correct in this case.

If we instead take the wide scope rational requirement to imply a wide scope ought, then we get a much more plausible result. The wide scope ought tells us that we are not as we ought to be in cases where we intend to eat Pot Noodles and believe that a necessary means to doing so is removing the lid, while at the same time not intending to remove the lid. In this particular instance, I can comply with the wide scope ought by not intending to eat the Pot Noodles.

The rational requirement of the instrumental principle can be expressed formally like this:

1. You ought (if you intend e and believe n to be a necessary means to e, then intend n)

    F1. RR[(Ie & BMne) $\rightarrow$ In]

Taking rational requirements to entail a wide scope *ought*, we can derive from this rational requirement the following normative requirement:

---

[71] See section 2.5.

2. You ought (if you intend *e* and believe *n* to be a necessary means to *e*, then intend *n*)

F2. O[(I*e* & BM*ne*) → I*n*]

According to standard deontic logic, the substitution of logical equivalents within the scope of an ought is allowed. By this rule, we can derive:

F2a. O~[(I*e* & BM*ne*) & ~I*n*]

F2a shows us that a person who has not complied with the instrumental principle is in the state where he intends some end and believes something to be a necessary means to it, without intending to do what he believes to be the necessary means. Such a person, if he fails to reason according to the instrumental principle, will not be coercible. You could threaten to have my car towed unless I moved it post-haste from your front garden. In making this threat, you would most likely be assuming that I intend to prevent having my car towed away and believe that you will do as you say. If I am instrumentally irrational, I shall not form the intention to remove my car, despite the fact that I have the intention and belief that you have attributed to me. Likewise in the case of the armed robber; convincing an instrumentally irrational victim, who intends to protect his family, that his family will be harmed unless the robber is given the combination to the safe will not effect the giving of the combination.

While rationality has provided us with a normative requirement that we act in

accordance with the instrumental principle, there is also a case for its being the case that an agent in this circumstance ought not to act in accordance with the instrumental principle. Because it is what would make the agent's life and that of his family's go as well as possible, he ought not to comply with the instrumental principle. We can formalise the non-compliance requirement like this:

3. You ought not (if you intend to *e* and believe *n* to be a necessary means to *e*, then intend *n*)

F3. O~[(I*e* & BM*ne*) → I*n*]

One difficulty with this solution is that F3 might be regarded as implausibly strong. The arrows in normative requirements are material conditionals and are not indicative of some kind of rational move or procedure. If we interpret F3 as a global requirement on our mental states, then not only will we be required by normativity not to have collections of beliefs and intentions that are relevant to the situation with the burglar, but we would also be required to rearrange all of our mental states such that there was no intention-belief-intention triad that meets the form proscribed by F3.

If this requirement is implausible, there is still a recourse in the armed robbery case. We can formulate a weakened version of the instrumental principle. This tells us that in the case of occurrent mental states, we ought to be *disposed* in a way such that if we intend an end and believe something to be the necessary means to that end, then we ought to intend what we believe to be the necessary means. Formalised, this is very close

to F2:

4. You ought to be disposed such that (if you intend *e* and believe *n* to be a necessary

means to *e*, then intend *n*)

      F4. OD[(I*e* & BM*ne*) → I*n*]

To thwart the armed robber, it would be sufficient not to have the disposition to intend

what one believes to be the necessary means to ends that one intends. Without that

disposition, you would be too unreliable to coerce. Reformulated this way, we arrive at:

5. You ought not to be disposed such that (if you intend *e* and believe *n* to be a

necessary means to *e*, then intend *n*)

      F5. O~D[(I*e* & BM*ne*) → I*n*]

Because of the strength of F2 and F3, for the rest of this section, I shall discuss

disposition formulation of the instrumental principle.

      We are now ready to see where a putative normative conflict arises from giving

the sort of normative status to rationality that has been given to it in this section. F4

expresses the requirement that we ought to comply with the instrumental principle:

F4. OD[(I*e* & BM*ne*) → I*n*]

On the other hand, F5 tells us that we ought not to comply with the instrumental principle:

F5. O~D[(Ie & BMne) → In]

From F4 and F5 we can derive F6:

F6. OD[(Ie & Bn) → In] & O~D[(Ie & Bn) → In]

The conjunction in F6 is a normative conflict: it is not possible both to comply with the disposition version of the instrumental principle and not to comply with it.

I shall say something in the next chapter about how we might think about coping with such normative conflicts, but first it is necessary to spend a little bit of time defending the view that there is really a conflicting set of requirements here. One way to try to account for this putative conflict is to argue that it is not the case that we ought not to comply with the instrumental principle, but rather that we ought to *cause* ourselves not to comply with it, making the negation range over an action, distinct and apart from the process of reasoning.

As for the cause account,[72] there are at least two important motivations for putting it forward. The first is that one might think that an agent cannot just fail to have his mental states conform to the requirements of rationality, but would need to cause himself

---

[72] There is a related discussion about causing oneself in section 1.4.

to do so. I shall address this consideration shortly. The second is that if it is rational to cause oneself to be irrational, the formal inconsistency appears to go away. I wish to consider this second motivation first.

Call this the cause yourself account. The burglar enters your home and you realise now that you must *cause yourself* to become irrational. Fortunately for you, there is a vial of temporary irrationality potion sitting on the table next to you, and you quickly consume it before the burglar can stop you. You consume the vial, causing yourself to be irrational.

In the situation where a vial of potion is available to you that will cause you to be disposed not to act on the instrumental principle, it looks plausible to think that you ought to cause yourself (by drinking the potion) to become irrational. On the cause yourself account, normativity does not directly require you both to comply with the instrumental principle and not to comply with instrumental principle. However, the cause yourself solution does eliminate the possibility of a normative conflict, and there is some reason to think that the cause yourself solution is not wholly accurate.

The cause yourself solution does not eliminate a putative normative conflict, because causing yourself to be irrational entails that you will be irrational. If you comply with the requirement that you cause yourself to be irrational, then that analytically entails that you will not comply with the requirement to be rational. Normativity still makes a demand on you that you cannot meet, and a normative conflict remains.

More importantly, the cause yourself solution is a dubious interpretation in the first place. If there were no reason for you not to comply with the instrumental principle, then there would be no reason in the case above for you to cause yourself not to comply

with the instrumental principle. The reasons for causing yourself not to comply with the instrumental principle are derived from those for not complying with the instrumental principle.

If rational requirements of the form discussed here entail that one ought to comply with the rational requirement, then putative normative conflicts follow. In chapter 4, the status of these possible normative conflicts will be discussed in more detail, but it is worth noting two points here.

The first point is that the putative conflict discussed here occurs between an ought deriving from a formal requirement on the relations among an agent's mental states and an ought deriving from the goodness, or rather badness, of the agent's mental states being in those relations. Putative conflicts of this kind could occur for any formal requirement on the relations among an agent's mental states.

The second point is that these putative conflicts do require serious consideration by normative theorists. Philosophers who feel that there are no conflicting all-things-considered oughts must either provide an account of how to resolve what they will regard as the apparent conflict in this example, or they must let go of the normative status of rational requirements.

3.3 Bad wiring

This section discusses putative conflicts between what an agent ought to do and what an agent ought to believe. The main example of this section concerns a human

agent. But, before presenting this example, I want to illustrate the possibility of this type of conflict in a slightly different way.

Consider a very poorly designed robot, 'Fred'. Fred is powered by a single battery. This battery has to power all of Fred's motor functions, his CPU, his memory, and all his input devices. This battery is capable of a certain maximum electrical output, m. Fred is physically unable to operate at a level that requires more than m. Because Fred has limited energy to expend, Fred is a pretty boring robot. Fred's lone output device is a robotic arm with a hand at the end. Fred has only one way of storing memory, too, and that is on an extraordinarily energy-inefficient RAM chip. In fact, this chip is so inefficient that storing and maintaining a single binary pair in his memory chip requires exactly the same amount of energy as moving his robotic little finger 1 cm.

Fred is fitted with a programme that tells him to store a binary pair b in his RAM and also to move his robotic little finger 1 cm. Unfortunately for Fred, the difference between M and the amount of energy it takes to keep Fred running in the most minimal possible way leaves him with just enough energy left over either to encode and maintain the binary pair b in his memory or to move his robotic little finger 1 cm, but not both. These two instructions to Fred cannot both be carried out; he can either perform the movement he has been told to perform or store and maintain the binary pair he has been instructed to store and maintain. What precisely would happen as a result of this energy shortage would depend on various precise details of Fred's construction that are not relevant here. What is relevant to note is that the one certain consequence of the energy shortage is that Fred cannot both move as he is required to move and to store and maintain memory as he is required to store and maintain by the programme.

Imagine now that Fred is a person and that moving his finger is an action, rather than just a movement, and that *b* is not a binary pair, but a belief. Furthermore, imagine the requirements put on him (to move his finger and to believe *b*) do not come from a programme, but from normativity### Fred ought to move his finger 1 cm and ought to believe *b*, but he does not have the ability to do both. Furthermore, imagine the requirements put on him (to move his finger and to believe *b*) do not from a programme, but from normativity### Fred ought to move his finger 1 cm and ought to believe *b*, but he does not have the ability to do both. Fred, in this case, would be stuck in a normative conflict, being unable both to do what he ought to do and to believe what he ought to believe.

This putative normative conflict has a somewhat different form to the one presented in 3.2. There the normative conflict arose because of the analytic impossibility of both complying with a requirement of rationality and causing oneself not to comply with that requirement of rationality. Here, the putative conflict arises as a result of the laws of physics. It is physically impossible for Fred both to do what he ought to do and to believe what he ought to believe.

Here is similar example, but one involving a person. Mary the anthropologist has been studying the difference between people who live on the Bristol side of the Avon Gorge and people who live on the Somerset side of the Avon Gorge. In an interview with a local newspaper, read on both sides of the gorge, Mary made some particularly disparaging remarks about both of her subjects of study. An angry mob of Bristolians has chased her onto the Brunel Bridge. Halfway across the bridge and running for her life, Mary sees another mob, this one of angry Somerset residents, approaching her. Both

groups are out for blood.

Mary quickly weighs up her options. She faces certain death if she runs one way or the other, or if she remains where she is. She faces a less certain fate### at least a non-zero chance of surviving, although not one with good prospects### if she jumps. Clearly, Mary ought to jump.

In addition to being a clever anthropologist, Mary is a former rescue worker and knows with great accuracy what her chances of surviving falls of different heights into water are. Furthermore, she is excellent at estimating heights. The bridge, as it happens, is at such a height that her chances of surviving would be greater than 0% and less than 1%, still better odds than she faces with the mob. However, these exceedingly dim prospects pose a real problem for Mary.

With the mobs approaching fast, Mary has only three seconds to jump before they reach her. The way her brain is hardwired, brain states that are tokens of the belief that doing q would be more than 99% certain to result in death always trigger a block on forming the intention to perform q for at least five seconds. If Mary comes to believe that jumping off the bridge will be more than 99% likely to result in her death, as the evidence available to her suggests, then she will not be able to jump off the bridge at all, as the angry mob will catch her before she can form the intention to do so. Believing what she ought to believe at time t, i.e. three seconds before the mob arrives, is not compatible with doing what she ought to do at time t, i.e. jump off the bridge. It is physically impossible for Mary, owing to her brain's hardwiring, to do what she ought to do and believe what she ought to believe. This is an inter-normative conflict between belief and action.

While this case is highly artificial in its construction, its overall form is not particularly unusual. A soldier might not be able to bring himself to storm an enemy position if he believed what the evidence suggested about his chances of surviving such a feat. In a basketball game, an outstanding free-throw shooter ought to concentrate all his attention on shooting the free-throw. This might absorb all his cognitive resources to the point where he is no longer capable of forming beliefs about the world around him. Even though he ought to believe that the crowd is cheering for him, being temporarily unaware of their existence, he does not believe that they are cheering.

Humans are beings with limited cognitive resources, whose brains are presumably not infinitely flexible in their wiring. As the examples of Fred and Mary suggest, there are times when it is physically impossible for people to believe what they ought to believe at a particular time and to do what they ought to do at the same time. This possibility means that people may periodically be subjected to putative inter-normative conflicts.

A common response to this sort of case is that there could not be any kind of inter-normative conflict. On the view of normative separatists, the normativity that applies to action and that normativity applies to belief are of fundamentally different, incomparable kinds. In chapters 1 and 2, I offered some support for the view that normative separatism cannot just be taken for granted, and in chapter 4 I shall discuss the matter again in more detail. Here I wish to introduce an additional argument, the argument from normative completeness.

Return for a moment to the analogy of a computer. Computers operate by running binary pairs through logic gates that produce new binary digits. Certain strings

of these pairs constitute *instructions*. Instructions tell the computer what to do, and the computer acts automatically on them.

Instructions can tell a computer to do all sorts of things, ranging from manipulation of external 'limbs' to the storing of data strings in memory. Computers can be instructed to change their 'beliefs', to operate external devices, and to read files for new instructions.

One thing that serial computers cannot do is execute inconsistent or contradictory instructions. If a computer is told to do two things that it cannot do, for example to make a certain portion of its monitor only red and only blue and not both, it typically will crash. One could give the computer contingency instructions. For example, one could programme it such that when instructed to produce an only blue and an only red patch in the same area of its monitor, then it will follow only the instruction to produce blue. Here we can say that the blue instruction defeats the red instruction in a particular instance. Failing this sort of defeasing instruction or some other default technique for resolution of the conflict, the computer will crash or fail to execute either instruction.

As argued above, it is relatively easy to conceive of situations in which a computer could not both do what it ought (was instructed) to do and believe what it ought to believe. Consider a causing case. The computer is instructed to believe *b*. Along with the instruction to believe *b*, it is instructed not to write on the only sector of the hard disk where *b* can be encoded, an 'action' restriction. The computer, of course, cannot comply with both of these instructions; the computer cannot both successfully write *b* and prevent itself from writing *b* on the same hard drive in the same sector.

In order to avoid having the computer crash and indeed to keep things running

smoothly, we might consider it a requirement of a well-designed operating system or programming language that these situations resolve; that there is some sort of rule for sorting out conflicts between instructions. Without such a rule, sufficiently complex machines might crash a great deal.

Normative separatists make a much lower demand on normativity than a computer scientist might make on an operating system. People are not required to have resolution techniques when normativity tells them to do something and believe something when they cannot do both. This is because they do not believe that there is a genuine conflict. The *ought* of belief and the *ought* of action are different and incomparable. Perhaps this is appropriate, but just as it would be unexpected for an operating system not to be able to deal with the sort of resource distribution problems we saw above in Fred (i.e. we would expect that the operating system would prioritise the action or the memory function), it seems mysterious as to why normativity, which if anything should be able to sort out our oughts, has no means of dealing with a large class of seemingly conflicting requirements. To put it another way, it is not hard to find situations where one cannot both do what one ought to do and believe what one ought to believe. There is no reason in principle why normativity could not be equipped to manage such conflicts. For all of the same reasons that we might expect a computer's operating system to resolve conflicting instructions, it seems that we should expect normativity to arbitrate between competing theoretical and practical requirements. It seems that a strong reason must be supplied for choosing to treat these apparent conflicts as incomparable. Just as the problem for Fred is an intelligible and interesting one, parallel problems in people should be taken seriously.

3.4 The power of positive thinking[73]

      In this section, a putative intra-normative conflict will be introduced, one between the requirements within the domain of theoretical normativity. In chapter 1 of this thesis, I argued that some of the main arguments that could be deployed to show that only epistemic or evidential reasons could count as normative reasons for belief did not deliver that conclusion at all. In this section, I am taking it for granted that there can be both evidential and non-evidential reasons for belief. In the next chapter, I shall discuss how evidential and non-evidential reasons (or to be more precise, epistemic and non-epistemic reasons) reasons for belief might interact. For the moment, however, I only wish to show the sorts of serious conflicts that can arise between what for now I shall call evidential and non-evidential oughts of belief.

      There is one second remaining on the clock, and the Lakers trail the Celtics by one point. Smith, a player for the Lakers, stands on the free-throw line, having been awarded two shots on a foul. It is the final game of the championship series. If Smith can make one free-throw, the Lakers will have a chance to win the game, and hence the championship, as the game will go into overtime. If he makes both free-throws, the Lakers will win the series outright. But, if he misses both free-throws, then the Celtics will win the championship. Smith is a career 50% free-throw shooter. As it happens, his performance at the free-throw line is linked directly to his confidence about his ability to make the upcoming free-throw.

---

[73] This term is borrowed from Harman's (1999) section heading about the same subject. I developed these arguments independently, but they are of very similar form to his.

To be precise, there is a function that maps Smith's free-throw percentage to his confidence, i.e. his belief about his chances of making each shot. As Smith gets less confident, he becomes more nervous and develops shakes proportionate to his anxiety. As a result, Smith's chance of making a free-throw is 5% below what he believes his chances of making a free-throw to be. Smith bases his belief about his chances of what his coaches tell him about his past performance.

Smith's coaches are well aware of just how much his confidence affects his shooting, so Smith has been systematically lied to throughout the season and told that he is a 55% free-throw shooter. Smith's coaches also kept it secret from him that he reliably shoots 5% below what he believes to be his chances of making a free-throw, something they have discovered through various drills. However, before this game, Smith became aware of both of his actual free-throw percentage and the relationship of his belief about his chances to his actual performance.

It is possible that Smith evidentially ought to believe that his chances of making a free-throw are 0%. That is because of what one might call naïve subtraction. Smith starts out believing that his chances are n%, but, knowing the function that maps his chances to his belief about his chances, he realises that his chances are n% minus 5. He then realises that his chances are really 5% below that, until he hits the lower limit at 0%. While the evidentialist has no other recourse for explaining why Smith ought to believe that his chances of making the free-throw are 0%, naïve subtraction is an unsatisfying account of why Smith ought epistemically to believe that he has a 0% chance of making his free-throws. One slightly more sophisticated consideration is that without a rule that says that there is a reason to believe the fixed point, theoretical rationality would be

unavoidably unstable in some circumstances. If we adopt the view that he ought to adopt the stable belief, which in this case is that he has a 0% chance of making his free-throws, he will miss all of his free throws. Zero is the lower limit for the range of free-throw percentages and represents the only stable belief at which Smith could arrive. For any other number, his belief will necessarily be unstable. (This more sophisticated reason for Smith to believe that his chances are 0% may not be available to a strict evidentialist for reasons discussed in section1.5.)

Now take a different example. Imagine sufferer of the dreaded disease *d*. *D* kills 90% of those whom it afflicts, and this figure is uniform across all groups of people except one. Those who believe that their chances of surviving *d* are better than 25% succumb only 85% of the time.

Bob contracts *d*. He reads the studies and sees that the best chance he could have of survival, even if he was extremely optimistic, would in fact be just 15%. Bob is very concerned with epistemic correctness. If he believes that his chances of surviving are better than 25%, then he knows that he actually only has a 15% chance of survival. He also knows that if he modifies his belief by that evidence, he will then have reduced his chances of survival to 10%. Bob decides that all the evidence in the end will point to his only having a 10% chance of survival. So, he adopts that belief. Bob, true to his epistemological self, believes what he ought to believe about his chances for survival.

Jane, on the other had, also contracts *d* and reads all the same studies. Jane, however, is an eternal optimist and believes that she is going to overcome *d* no matter what, as she holds the erroneous belief that enough determination will let you do virtually anything. Jane is in the epistemic wrong, not believing what she epistemically ought to

129

believe. But, as long as goodness based reasons for believing are allowed, she does in

another sense believe precisely what she ought to believe. By being so optimistic about

her chances of survival, she has given herself a 5% improvement in her chances of

surviving *d*. Jane also believes what she ought to believe, except that 'ought' this time is

qualified as being non-epistemic.

Likewise with Smith. If he forms a belief that his chances of making that free-

throw are certain, then he will believe what he ought, non-epistemically, to believe, as he

will then have a 95% chance of making the shot. On the other hand, he will believe what

he ought to believe evidentially if he arrives at the conclusion that he is certain to miss.

The *d* sufferers and Bob are all under the sway of conflicting oughts, where what

the evidential or epistemic considerations tell you that you ought to believe is quite

different to what the goodness based considerations tell you. Provided that the door

cannot be shut on non-evidential reasons for belief, there will always be the possibility of

conflict between what one ought epistemically to believe and what one ought non-

epistemically to believe.

Conclusion

In this chapter I have argued that there are several normative conflicts, or apparent

conflicts, that theories of normativity must be prepared to address. Theories that are

committed to the view that normative conflicts are impermissible must account for and

resolve the apparent inter-normative and intra-normative conflicts presented in this

chapter. Theories that admit of the possibility of normative conflicts must provide a

principled account of which of these are conflicts in normativity and which are not. Theories that do not admit of normative conflicts must explain why these conflicts are only apparent.

In chapter 4, I shall discuss further the ramifications of these putative conflicts and shall offer a tentative account of how theorists who do not wish to admit conflicts into normativity may seek to resolve at least some of them.

Chapter IV: What to Do about Putative Normative Conflicts

Introduction

Chapter 3 discussed putative normative conflicts. These conflicts occurred when it was not possible for an agent to comply with two or more requirements of normativity. While chapter 3 made the case that these putative normative conflicts could arise in a variety of circumstances, it did not discuss the broader ramifications of putative normative conflicts for theories of normativity. This chapter examines some of the ramifications of putative normative conflicts for theories of normativity, and it also looks at the ways in which different types of normative theories might address putative normative conflicts.

This chapter is divided into two main parts. The first part is comprised of sections 4.1 and 4.2, which discuss *normative architectures*. A normative architecture is a structure for a normative theory that accounts the sources of normativity, e.g. truth and goodness, available to the theory and how they relate to the domains of normativity, or types of reasons and oughts, contained within the theory. The analysis of normative architectures is provided is done with three aims in mind. The first aim is to provide a general framework for considering how domains and sources of normativity might be related to each other. Looking at this relation allows for an improved understanding of normative separatism, as normative architectures provide a framework for an analysis of normative separatism. The second aim is to help make clearer what sort of meta-normative commitments are more and less friendly to normative theorists who accept,

and that do not accept, normative conflicts. Finally, the discussion of normative

architectures helps lay the groundwork for the second section of the chapter.

This second part begins with section 4.3. It discusses some approaches to

addressing putative normative conflicts. Using a putative conflict within theoretical

normativity as a paradigm, the notion of *normative defeasing* will be introduced as a

possible approach.

Putative normative conflicts can occur under different circumstances, as was

discussed in chapter 3. Some derive from competing sources of normativity within one

domain### e.g. theoretical normativity- and others from the competing demands of

different domains of normativity### e.g. when one cannot both do what one ought to do

and believe what one ought to believe. As an instance of the first sort of conflict, I

argued that sometimes within theoretical normativity, what there is most reason to

believe based on the evidence may be inconsistent with what there is most reason to

believe owing to non-evidential considerations. I also provided examples of the second

sort of conflict; one case involved having insufficient mental resources to both do what

one ought to do and believe what one ought to believe. Another version of the second

sort of putative conflict that was discussed occurs when one is required to cause oneself

not to believe something that one ought to believe. Putative conflicts of the second sort

also come from the competing normative requirements arising from rational requirements

(if rational requirements entail normative requirements) and from the normative

requirements that derive from the benefits of not complying with rational requirements.

As an example of this type of conflict, I discussed instances in which it might be best for

us not to be in compliance with the instrumental principle because things will go much

better for us if we are not.

There are two distinct routes to addressing these putative normative conflicts, because they can be addressed at two different levels: the meta-normative level and the first-order level. At the meta-normative level, one might adopt normative architectures that do or do not allow for such conflicts. If a normative architecture does not allow for normative conflicts, then one must provide first-order solutions to the conflicts. If it does allow for normative conflicts, then no first-order solution is required. These two levels require some fleshing out.

Let us begin with the meta-level. A complete theory of the architecture or logical structure of normativity would have to tell us what sorts of things there are reasons for, such as beliefs, desires, actions, intentions, emotions, etc. It would also have to tell us something about where those reasons came from. the source of normativity### evidence, benefit, and so on. Does each different type of thing for which there are reasons or oughts have a different source of normativity? Do these sources of normativity for one domain of normativity also serve as sources of normativity for other ones? We would also need to know whether or not (and if so, how) the reasons and oughts that arose in each domain of normativity (the practical, the theoretical, and so on) could be weighed against or otherwise compared with the reasons and oughts from other domains of normativity.

Looking at different possible architectures is important in accepting or denying a position that has importance in the discussion about possible normative conflicts. That position is normative separatism. For example, normative separatists believe that when the ought of theoretical normativity and the ought of practical normativity conflict, there

is no further question as to whether one ought to comply with the theoretical or practical ought. Practical ought and theoretical ought are basic concepts; there is not a more basic ought that subsumes them. Normative separatists need not deny that practical and theoretical normativity have various internal structural affinities with one another, but they are not logically related in such a way that the demands of different domains of normativity are comparable.

We can attempt to address at least some possible normative conflicts at the meta-normative level. Adopting normative separatism puts conflicts between theoretical and practical normativity into a particular form. As they both are by definition basic (that is, there is no more fundamental 'just plain normativity' underlying them) and incomparable, then normative separatism tells us that at times agents are unable at times to satisfy the demands of both.

There is a limit to how far one can go towards resolving putative normative conflicts at the meta-normative level, however. This becomes apparent as in some of the architectures putative normative conflicts are in principle resolvable, but nothing about the architecture itself provides direction on how to resolve them.

Looking at first order ways of resolving putative normative conflicts, this chapter discusses ways in which they might be addressed within normative architectures where such conflicts are meaningful. The main part of this discussion concerns the normative defeasing relation, which, I argue, looks like a promising way of accounting for why certain apparent conflicts look problematic but are in fact resolvable.

Before continuing with the analysis of possible normative architectures, it is important to note that these architectures may be unsatisfactory in many ways. I do not

mean, or just mean, that some of these architectures are not promising as far as normative architectures go, but that the concept of a normative architecture itself may be flawed.

One of the concepts involved in describing normative architectures is that of a domain of normativity. A domain of normativity corresponds to a type of reason. For example, the domain of theoretical normativity is that part of normativity that is concerned with belief, and the domain of practical normativity is that which is concerned with actions and related mental states. This concept is not particularly problematic. Domains of normativity are identified by what sorts of things the reasons and oughts in them are reasons and oughts for.

The other important concept that is used in defining a normative architecture is that of a source of normativity. This concept is problematic. While a partial analysis of it is given in section 4.1, I shall highlight why it is problematic here.

A source of normativity is something like that in virtue of which a reason relation holds, or a particular kind of explanation of why a fact is a reason of a particular type. For example, the fact that there is a seminar today is a reason for me that I believe it is not the weekend. This fact is a reason for this belief because this fact is evidence for this belief. The source of normativity for the reason is evidence, or perhaps truth.

Sometimes it might be difficult to say why something is a reason for something else, even though it no doubt is. Suppose Loretta's window sill is unadorned and has no plant on it. The fact that Loretta's window sill does not have a plant on it is a reason for you that you give her a plant. What is the source of normativity? It might be benefit, as Loretta will be happier if she has a plant on her window sill. It could be aesthetics, as the window sill will not look very nice without a plant on it. Or, perhaps, there is nothing

much more to say about why the fact that Loretta's window sill does not have a plant on it is a reason for you that you give her a plant. At least, I am not certain what is the correct thing to say in this case.

Although I find the concept of a source of normativity to be somewhat unclear, I have pressed ahead and used it. The reason for using it is that it is a concept that seems to me to play an important role in the way some philosophers think about normativity.[74] The underlying thought is, as far as I can tell, that certain types of reason only issue from certain sources of normativity. Evidentialists think, for example, that reasons for belief only issue from evidence. Strict consequentialists about reasons for action believe that reasons for action only issue from features of the states of affairs that will (or are expected to) result from the action in question.

It may that it is best to do away with sources of normativity and, consequently, normative architectures. Nevertheless, as the concept of a source of normativity, or one close to it, plays an important role in shaping the philosophical debate, it seems best to work with it, at least to see what follows from accepting there is such a concept.

4.1 Possible normative architectures 1: basic and derivative normativity

In order to begin exploring normative conflicts, two notions must be introduced. They are the notions of *basic normativity* and *derivative normativity*. Value theorists distinguish between intrinsic and instrumental value, the notion being that some things

---

[74] See for example Hieronymi (typescript), Railton (1994), Railton (1997) Skorupski (forthcoming) and (1999), and Wedgwood (2002*b*).

are valuable in and of themselves and others are valuable in virtue of their instrumental relationship to something that is itself intrinsically valuable. I shall offer a similar distinction for normativity, one between basic and derivative normativity. Before I do that, however, I must add a caveat. Since the notion of intrinsic value was first discussed extensively by G.E. Moore,[75] there has been a good deal of controversy concerning whether or not there are other sorts of not-derived value, perhaps relational or extrinsic non-instrumental value. This suggests that a better way to discuss value may be in terms of basic value and non-basic value. Basic value would include intrinsic value, extrinsic value, and any other sort of value that did not itself rely on some other sort of value. Derived value is that which derives its value from basic or other derived values.

In chapter 1 of this thesis, the structure of reasons was discussed.[76] A reason is a fact that stands in the reason relation to an agent and a proposition concerning an action, belief, feeling, or whatever else there can be reasons for.[77] Not yet discussed in any detail is the logical structure of ought. Because the possible normative conflicts set out in chapter 3 were characterised in terms of oughts, it is necessary to say something about the logical structure of ought. Ought is an operator that operates on propositions. We can analyse any ought sentence, at least when ought is serving its role as, to borrow Bernard Williams's phrase, the central normative operator, as saying 'Ought p', where p is some proposition.

Initially it might look unlikely that ought governs a proposition, as the normal English usage of 'ought' has it taking an infinitive. 'Ought that' is not a normal locution

---

[75] See especially *Principia Ethica* and Moore's lesser known essay, 'Intrinsic Value'.

[76] See section 1.1.

[77] The three place relation is not exhaustive, as there may be further places for a time, circumstance, degree, etc. See footnote 3 in chapter 1 for references relating to this

in English, and we normally expect operators that take propositions to be followed by a that rather than an infinitive. There are locutions in English, which in many contexts mean the same thing as ought sentences do, that take 'that'. As two examples, one may say, 'It is fitting that he goes the store' or 'It is meet that Achilles honours his friends.' More importantly, the intuition that infinitives are not expressing propositions can be a mistake, as infinitives contain tacit subjects. There need be nothing conceptually different between the locutions 'He ought that he goes to the shop' and 'He ought to go to the shop', since 'to go to the shop' in fact has a subject, which is 'he'. However, there is an advantage to using the non-grammatical 'ought that' over the grammatical 'ought to', and that is the ease with which one can make the tacit subject of the sentence that ought governs explicit. There are other good reasons for using 'ought that' in this context, the primary one being that it allows for a convenient way of expressing an ought governing a conditional sentence, such as 'He ought that if he goes to the store, then he buys some eggs'. In normal English usage, an ought governing a conditional is attached to the consequent of the sentence, 'If he goes to the store, the he ought to buy some eggs'. Unfortunately, this is the identical locution used in English to express the case where the ought is logically attached to the consequent. Ordinary English has no grammatical means for distinguishing between oughts that govern only the consequent and those that govern whole conditionals.

It may seem at first that we can understand the sentence 'I ought to go to the shop' as logically expressing: 'ought that I go to the shop'. That expression is incomplete, because the ought operator needs a subject. To provide a more accurate ought sentence, we can use what I shall call an *O-form sentence*: 'I ought that I go to the

---

matter.

shop'.

In the O-form sentence 'I ought that I go to the shop', the subject of the ought and the subject of the proposition that it governs are the same. Nevertheless, the subjects are logically distinct entities, and the O-form sentence picks this up. It is in principle possible for the subject of the ought to be different from the subject of the proposition that it governs, although such cases will not be discussed here.

Although oughts and reasons have different logical structures, they are similar in three respects that are relevant here. First, they both index to an agent### for example A ought that A φ, and f is a reason for A that A φ. Second, we can give a type for both oughts and reasons by what they are oughts or reasons for. That is to say, a reason is a reason for belief (a theoretical reason), just if the relation is of the form fact f is a reason for agent A that A believe b and an ought is a 'belief ought' just if the O-form sentence is of the form A ought that A believes b. In this chapter, when the expression 'type of ought' is used, it refers to whether the O-form sentence is an action sentence, a belief sentence, a feeling sentence, etc. Following the convention established in chapter 1, I shall use 'type of reason' in the same manner. Finally, both oughts and reasons govern propositions. As explained in section 1.1 of this thesis, I shall sometimes discuss reasons for action or for belief rather than reasons for propositions concerning actions or propositions concerning beliefs. I do this because the terminology of propositions concerning actions and propositions concerning beliefs quickly becomes cumbersome. I believe that this economy will not prove too confusing.

Before continuing, it is important to provide an analysis of the concept of a *source* of normativity. It is perhaps confusing and misleading to suggest that normativity has a

source. The use of the word 'source' might wrongly be interpreted to imply that normativity flows from some interesting metaphysical construction in the way light does from the sun. A better analysis will appeal instead to a particular sort of explanation of why an ought or a reason holds.

Assume that the fact that Mary is drowning is a reason for James that he save her. One can ask what sort of reason this is. It is a reason for an action proposition, that James save Mary. One can also ask why the fact that Mary is drowning is a reason for James that he save her. One might answer by saying that in these circumstances, Mary would benefit from being saved by James. That it benefits Mary in these circumstances is the explanation of why the fact that Mary is drowning is a reason for James that he save her.

Consider, as well, a case involving a reason for belief. The fact that it is Tuesday is a reason for me that I believe that yesterday was Monday. Here the explanation of why this fact is a reason for me that I believe yesterday was Monday is that the fact that it is Tuesday is evidence for the fact that yesterday was Monday.

In both of these examples, there is an explanation available as to why some fact is a reason for something. The sort of explanation that I have in mind is some feature of what makes the reason relation hold between the fact, the agent, and what the reason is for. In the case of the action above, it is that it is beneficial to Mary (or just generally) that James save her when Mary is drowning in this instance. In the belief example, it is that the fact is evidence for the contents of the belief.

The notion of explanation here is not meant to imply that all explanations of reasons will be quite as simple as benefit or evidence. There may be an unlimited

number of normative sources or very few. Radical pragmatists, for example, might think that benefit is the only source of normativity for any type of reason. A very radical pluralist might think that each reason has a slightly different source. The possible relations of sources to reasons will be discussed below.

We can move on to discuss basic and derivative normativity. Basic normativity is normativity that does not derive from its object's role as a means to, in promoting, or of as a component of some other normative end. Derivative normativity is any normativity that is not basic.

As an example, consider a simple theory of normativity that tells you that you have reason to do whatever will make you happy. Looking at a case involving reasons, the fact that eating sweets will make you happy is a reason for you to eat sweets. The normativity of that reason is basic. Eating sweets is one of the things that makes you happy. Saving money, on the other hand, does not make you happy. However, in order to buy sweets, you must have saved some money. So, there is a derivative reason to save your money. The fact that saving money will allow you to buy sweets (which in turn will make you happy) is a reason for you to save your money.

4.2 Possible normative architectures 2: combinations of basic normativity and types of reasons

This section discusses the ways in which sources of basic normativity can combine with different types of reasons and oughts. In this discussion, some more

terminology is required. At the risk of adding another use to the already overlong list of monisms and pluralisms, this chapter will make use of the term *normative monism* and *normative pluralism*. These are claims about the number of sources of basic normativity. Normative monism states that there is only one source of basic normativity. So, a radical pragmatist who thought only goodness provided reasons of any type would be a normative monist. Normative pluralism is the view that there is more than one source of basic normativity.

The other conceptual axis is *reasons* (or oughts) *specialism* and *reasons* (or oughts) *generalism*. This distinction requires some care in specifying, as there are some species of reasons generalism that look like reasons specialism.

Reasons specialism is a view about the conceptual connection between sources of normativity and types of reason. Reasons specialism holds that for each type of reason ### e.g. for action propositions, belief propositions, and so on### the source or sources of normativity from which it issues have an appropriate conceptual link with that type of reason. For example, evidentialists may think that it is a conceptual truth that reasons for belief are always evidential reasons and never reasons deriving from the goodness of believing something. The reasons specialist holds that this is a conceptual truth about reasons for belief. By way of contrast, a non-specialist might still hold that there are no goodness-derived reasons for belief, but that this is not a conceptual necessity. Reasons generalism is the negation of reasons specialism. It holds that it is conceptually possible for any source of normativity to issue any type of reason, even if it is the case that there are no actual instances of a particular type of reasoning issuing from a particular source of normativity. This distinction is more easily discussed in context and is elaborated on

more fully in 4.2.2 below.

One preliminary difficulty with the distinction is that there are some reasons generalist views that will not have certain basic sources of normativity applying to all types of reasons, and there could be a reasons specialist view in which all sources of normativity are bases for all types of reasons. As an instance of the former, it seems possible that while there could be goodness-based reasons for belief, it would be very strange if there were any evidence-based basic reasons for action, other than perhaps for speech acts. Whether such a position is genuinely a reasons generalist one depends on why there is not a match between a particular basic source of normativity and a type of reasons. This matter requires more explanation, which will be provided in 4.2.2.

4.2.1 Architecture 1: normative monism and reasons generalism

Having presented these two axes, it is now possible to discuss the various possible architectures of normativity. The first architecture is normative monism and reasons generalism (NMRG). The normative monism part says that there is only one source of normativity. The reasons generalism part says that all sources of normativity in principle can apply to all types of reasons. An example of a theory that conforms to NMRG is radical pragmatism. A radical pragmatist thinks that the only reason for anything, be it a belief, an action, or a feeling, is that the thing in question will lead to an increase in overall goodness.[78]

NMRG has powerful resources for resolving putative inter-normative conflicts,

_____

[78] This form of radical pragmatism is a rarely held position. The major work on the subject is Stich (1993).

e.g. putative conflicts between the requirements of theoretical reason and those of practical reason. To see why, consider what the normative world would be like if the only source of normativity were goodness. A fact would be a reason to do or believe something for an agent only when that fact made it so that the performance of the action or the holding of the belief would increase total goodness. Furthermore, we could assign to each reason a weight based on how much better things would be if the action proposition or belief proposition for which there was a reason obtained.

In this circumstance, figuring out what we ought to do or what there is most reason to do would be a matter of weighing reasons deriving from the same source. On this very simple theory, how much reason there is depends entirely on how good the outcome is of doing or believing what there is a reason to do or believe. One *really* ought to believe what there is most reason to believe if that produces a better result than doing what there is most reason to do, and *vice versa*.

This is not to say that NMRG does not admit any troubling cases. Presumably as there is sometimes thought to be incommensurability, or more importantly, incomparability in value,[79] there might also be incomparability among reasons deriving from a single source of normativity. This incomparability might occur if the source of normativity were value, and the reasons that issued from it kept the same logical structure as the value. Instances, were there any, of incomparable values would issue incomparable reasons. But, that said, there would at least be no mystery about which feature to compare when evaluating different types of reasons, since all types of reasons would have the same source of normativity. One would weigh up all reasons with respect to their single source of normativity. If all reasons came from goodness, then we just

145

have to look at how to weigh up goodness-based reasons.

4.2.2 Architecture 2: normative monism and reasons specialism

The second possible normative architecture is normative monism and reasons specialism (NMRS). The normative monism part says that there is only one source of normativity and the reasons specialism part says that each type of reason is conceptually connected to its sources of normativity.

NMRS is a less plausible normative architecture, because it may rule out the possibility of there being certain types of reasons. Consider a form of NMRS in which the sole source of normativity was evidence. While the link between evidence and belief has much plausibility, it is unclear what the conceptual relationship would be between there being evidence for something and one's having reason to do it. Of course, this problem applies to NMRG, too, but not all sources of normativity will prove equally problematic for NMRG as they do for NMRS. Looking at why some cases are more difficult for NMRS than they are for NMRG provides an opportunity to elucidate the distinction between reasons generalism and reasons specialism.

Reasons generalism and reasons specialism distinguish between two different conceptual connections between sources of normativity and types of reasons. Reasons specialism says that each type of reason, if there are genuinely reasons of that type, has a source or sources of normativity that have an appropriate conceptual connection to it.[80]

---

[79] See Chang (1997).
[80] The fitting attitude view is a good example of this. See Scanlon (1998), Ewing (1959), Rabinowicz and Rønnow-Rasmussen (forthcoming), and Hieronymi (unpublished typescript).

The constraints on these connections could vary from theory to theory, but the important contrast is with reasons generalism, which in principle allows any source of normativity to be a basis for any type of reasons. It may be discovered *a posteriori* that some sources of normativity in fact generate no reasons of a certain type, but this is not because reasons generalism excludes that source of normativity from being a basis for a particular type of reason *a priori*.

One view that NMRG can accommodate easily is radical pragmatism. Radical pragmatism says that there is a reason to do or believe whatever will be beneficial to do or believe. In an NMRG architecture, radical pragmatism can be stipulated. While one might need to make the case for pragmatism, one does not need to show that there is some particular type of connection between believing and goodness. In an NMRS architecture, however, it is not good enough just to show that goodness is the source of normativity. It also must be shown that goodness connects in an appropriate way with belief so that there can be goodness-derived reasons for belief. If this connection cannot be shown in an NMRS architecture, then there will be no reasons for belief in that case.

What is also interesting to note is that NMRS and NMRG can, with the same source of normativity and the same reason types, both yield the result that there are no reasons of certain types. This will be clearer with an example.

A normative monist could claim that the only source of normativity is evidence. If that normative monist were a reasons specialist, then by definition evidence could only be a basis for a type of reason that connected in the conceptually correct way with it. Most plausibly, reasons for belief would be said to connect appropriately to evidence, owing to the connection between truth and belief, while there is no such clear relationship

147

between truth and action. In this NMRS theory, there would be no reasons for action, because evidence lacks the kind of conceptual connection with action that it has with belief.

One might get the result that there was only one type of reason, too, in an NMRG theory that took the only source of normativity to be evidence and held that there were only two types of reasons, reasons for belief and reasons for action. Evidence serves as a basis for reasons for belief in an NMRG theory, as in an NMRG theory the source of normativity can in principle be a basis for any type of reason. However, an NMRG theorist might discover *a posteriori* that there are no evidential reasons for action. So, although there is nothing in the structure of the NMRG theory to rule out the possibility that evidence could serve as a reason for an action, there are in fact no evidential reasons for action. Although the NMRG theory and the NMRS theory yield the same result with these inputs### that the only source of normativity is evidence and that there are only reasons for belief### they yield that result for very different reasons.

4.2.3 Architecture 3: normative pluralism and reasons generalism

.

Normative pluralism is the view that there is more than one source of normativity. Normative pluralism and reasons generalism (NPRG) is the view that there is more than once source of normativity and that all sources of normativity can in principle serve as bases for all types of reasons. Consider the normative universe in which there are two sources of normativity, goodness and evidence, and also two types of reasons, those for action and those for belief. If this normative universe is of the NPRG structure, then it is

open conceptually for there to be evidential reasons to act[81] and also to believe, as well as goodness based reasons to act or believe.

It is not entirely clear how strong NPRG's logical resources are for dealing with competing normative claims among different types of reasons. Initially it may look as though NPRG has powerful resources for resolving putative normative conflicts, as it shares the main advantage of NMRG, namely that all types of normativity apply to all reasons. This is probably an optimistic assessment of NPRG. To see why, we can look first at an example of a theory compatible with NMRG, radical pragmatism. Radical pragmatists hold that goodness is the only source of normativity. Possible inter-normative conflicts are still just possible conflicts between good-based reasons. Such conflicts are addressable by whatever means are available for weighing only good-based reasons, even when weighing up reasons to believe against those to act. Of course, there may still be problems in comparing goodness based reasons, but these problems do not arise from confusion about how to weigh up evidential reasons for belief against goodness-based reasons for action, for example.

It is true that in NPRG, that we might think that all reasons that come from goodness have enough in common to find ways of coping with putative conflicts between goodness-based reasons, whether they are reasons for action or belief. However, it is not at all clear what to do with putative normative conflicts, whether they be within belief, within action, or between the two, that involve reasons with different sources of

---

[81] It is not entirely clear what an evidential reason to act would be, and I am not asserting here that there are such reasons. It is merely not excluded *a priori* by the structure of the theory that there are such reasons. If there are any evidential reasons to act, I suspect that they will be reasons for speech acts. Perhaps the fact that there is evidence that it is the 21st of the month is a reason to reply 'The 21st' when asked 'What is the date?'

normativity. We might expect a theory of goodness or a theory of evidence to be structured in such a way that they can manage their own problems, so to speak, but it is not clear why a theory of either would have much to say about how it relates in a non-derivative way to the other.

To put that point another way, a theory of goodness should be structured such that between any two states of affairs, either one is better than the other, they are equally good, or, perhaps, the value of one state of affairs cannot be compared to the other. Likewise, within a theory of evidence it might be expected that two facts can be evaluated as stronger or weaker evidence for (or against) a particular hypothesis, or as equally strong. Theories of evidence might even admit incommensurability. But what you would not expect from a theory of goodness is a method of how to measure goodness against evidence. You also would not expect a theory of evidence to provide a method for measuring evidence against goodness.

NPRG theory, on the other hand, has to find a method for weighing or comparing goodness-based reasons and evidence based reasons, if it is to provide a method for resolving putative normative conflicts involving reasons with both normative sources. There are at least two roads that NPRG might go in attempting to deal with possible conflicts where the reasons involved have different normative sources. The first looks the more promising one.

The first road arrives at the view that there is nothing unexpected about the idea that a theory of normativity tells us how to relate the plurality of normative sources. After all, a theory of value should tell us how a plurality of goods are related, even if it only says that they are incommensurable, and a theory of evidence might be expected to

tell us how a plurality of different sorts of considerations affect the likelihood of a proposition's being true. The theory of normativity, on this view, is a theory that tells us how all the sorts of normative sources relate to each other, and in that way is like any other theory that has a plurality of ontologically basic entities that have to interact with each other.

The second road takes us in the opposite direction and arrives at the view that there is something wrong with expecting a theory of normativity to do as well at resolving the problem of comparing reasons with different normative sources as a theory of value does at addressing how to compare different values. On this view, it seems that for different sources of normativity to be comparable, there has to be something in virtue of which they can be compared. I do not mean this as the start of a 'third man' argument, but rather that the very idea that we *could* compare two different sources of normativity suggests that either one is derivative of the other, and hence not basic, or that they are both derivative of a third source of normativity. This road leads to the conclusion that there are genuine normative conflicts.

The burden of proof appears to be on the second view. Someone who takes that view would need to show why it is less plausible that a theory of normativity could have a structure for inter-normative comparisons when there are a plurality of basic normative sources than that a theory of goodness could have a structure for comparisons among a plurality of values.

4.2.4 Architecture 4: normative pluralism and reasons specialism

The final normative framework discussed here is the one that seems to be most often taken for granted in the literature. This is normative pluralism and reasons specialism (NPRS). Normative pluralism says that there are a plurality of sources of normativity, and reasons specialism tells us that the sources of normativity must have the right sort of conceptual link with a type of reason in order to serve as a basis for that type of reason. The position I have earlier identified as *normative separatism* is a variety of NPRS, as it says that in principle only a certain source of normativity, evidence, has the right kind of conceptual link with a particular type of reason, reasons for belief, and that there are no other sorts of reasons for belief.

A caveat is required in saying that normative separatism is a species of NPRS. It is possible to have a mixed normative architecture, one in which some types of reasons must the appropriate special connection with their sources of normativity, while other types of reasons do not require this special connection. A normative architecture might say that as far as reasons for belief go, reasons specialism is true. On the other hand, as far as other types of reasons are concerned, there need not be any special conceptual connection between the reasons and the source of normativity from which they issue.

On NPRS views, the spheres of normativity are distinct. To speak of theoretical normativity and to speak of practical normativity is to speak of two unrelatable domains. This claim may sound strong, given that it is possible in principle to have the same source of normativity be a basis for different types of reasons. The unrelatability comes from the requirement that a source of normativity have a particular conceptual connection with the type of reason it serves as a basis for. Benefit, for example, might be a source of normativity for both action and feeling. That doing *x* is beneficial is a reason to do it. It

may also be a reason to feel good about yourself for doing it. However, the conceptual connections between acting and benefit and between feeling good about oneself and benefit are likely to be quite different, as acting concerns promoting benefit and feeling good about oneself involves responding to the fact that someone or something has been benefited.

Given this situation, NPRS lacks the logical resources to say anything at all about putative conflicts between our reasons for actions and our reasons for belief, because they are wholly different sorts of things. Asking what it is that one ought to do when one cannot act as one ought to act while also believing what one ought to believe is, by NPRS, a meaningless question.

The upshot of NPRS is that there are no inter-normative conflicts, because there is no unified concept of normativity that subsumes both theoretical and practical normativity. This is the view of normative separatism. I have argued in section 3.3 that it is an overly undemanding view of normativity that it not address cases in which one cannot do what one ought to do and believe what one ought to believe.

4.2.5 A final comment on normative architectures

There is an important issue that has not yet been discussed, one which relates to reasons specialism. This is a discussion of what constitutes an appropriate conceptual connection. Reasons specialism requires there to be an appropriate conceptual connection between a type of reason and the source or sources of normativity from which it derives. There are many possible conceptual connections, and merely identifying that

reasons specialism requires the appropriate one is not very illuminating.

It is not clear how informative the account of an appropriate conceptual connection can be. Consider reasons for thick pro-attitudes. Admiring is a good example. The source of normativity that most obviously has a close conceptual connection to admiring is admirability. We might be able to pick out some common features of things that are admirable, but 'admirable' itself is a normative concept### it denotes the sort of thing that one ought to admire### that does not obviously reduce to a cluster of other concepts. It may very well be a case by case question as to whether or not a particular thing is admirable, and there may be no possibility of giving general principles.

As an example of a thin normative concept, it might generally be thought that the fact that an action will produce some benefit is a reason to do it. Here there is some degree of mystery as to why benefit has the right kind of conceptual connection with actions *per se*. In general, that something would produce beneficial results seems to count in favour of it. It is unclear why, one the one hand, when beneficial outcomes are the result of an action, there is an appropriate conceptual connection between benefit and one's reason to act, while, on the other hand, when the beneficial outcomes are the result having a belief or a feeling, that connection would be absent between one's reasons and benefit. A major task for the normative separatist or any reasons specialist is to explain two things. First, what these appropriate conceptual connections are, and, second, why they are significant.

4.3 What to do about normative conflicts

The previous section discussed basic types of normative architectures, meta-normativity, if you like. This section looks at normative conflicts and examines possible first-order resolutions to them. Although there are a variety of putative normative conflicts, this section will examine one sort in particular. That is the conflict between epistemic and non-epistemic reasons for belief. It is a supposition for sake of argument in this section that epistemic and non-epistemic reasons for belief can be compared. If they cannot, then there are first order techniques for resolving conflicts between them.

The paradigm conflict in this section is one between epistemic reasons for belief and reasons for belief that arise out of the goodness of having that belief. It is worth saying a few words about this conflict, in particular by way of a brief reminder about conflicting reasons. It will be also be helpful to say something about why epistemic and non-epistemic reasons, rather than evidential and non-evidential reasons, are being compared.

Evidential reasons are a subset of epistemic reasons. The class of epistemic reasons includes### in addition to evidential reasons### reasons of the sort discussed in section 1.5. For example, the fact that you will be right if and only if you believe something is (or may be) an epistemic reason to believe it. However, it is not an evidential reasons, as that you will be right if and only if you believe something is not evidence that it will be true.

The example of the numbers game in section 1.5 showed the difficulty that evidentialism has with certain epistemic scenarios. Because of the issues raised in the numbers game, evidentialism is not a promising theory of reasons for belief. In addition,

some of the more interesting putative conflicts between different types of reasons for belief may involve epistemic, but non-evidential, reasons for belief. Because of these considerations, I have considered how epistemic and non-epistemic reasons can be compared, rather than evidential and non-evidential ones.

Reasons in this thesis have been assumed to be weaker than oughts. The fact that there is a reason for me to do something does not entail that I ought to do it. There is a reason for me not to pay my taxes### filling out the forms takes a lot of time, taxes deprive me of some of my income. On the other hand, paying my taxes helps to keep all the valuable services that the government provides going, paying them is a civic duty, it is a good example for others, and I benefit greatly from being in a society where people by and large pay their taxes. So, I most certainly ought to pay my taxes, even though there is a reason for me not to pay them.

When it is the case that there is a reason for me to do something and also the case that there is another reason for me not to do it, that does not count as a proper normative conflict. Reasons can weighed or compared in other ways, and it is the very nature of normativity that there are often reasons for and against the same thing, or that there are reasons that cannot all be complied with. The putative conflict in the forthcoming example is a conflict between what one ought to believe, *qua* epistemic reasons for belief, and what one ought to believe, *qua* non-epistemic reasons for belief. On the other hand, when I talk about evidential reasons to believe conflicting with reasons to believe that derive from the goodness of having that belief, it may not appear that there is any meaningful conflict, as we often comply with all of our reasons . In a sense, this is the very point that I want to consider, that putative normative conflicts are apparent rather

than real normative conflict. Indeed the purpose of this section of the chapter is to offer a structure for understanding why at least some of what appear to be normative conflicts are not conflicts at all.

This discussion is carried out within theoretical reason, as the difference between epistemic and non-epistemic reasons for belief is fairly intuitive. We can talk about we ought to believe *qua* epistemic considerations, and what we ought to believe *qua* non-epistemic considerations### in this case goodness. As far as goodness is concerned, I ought to believe *b*, whereas so far as truth is concerned, I ought not to believe *b*. The two different sources of normativity involved pull us in these instances in opposite directions. Traditional epistemology, insofar as it is concerned with theoretical reasons, is only concerned with epistemic reasons. The pragmatism of the Stich variety is only concerned with the goodness of our beliefs. Here the goal is to consider both sources in tandem.

I think it is likely that the general form of the model developed using conflicts within theoretical reason might be extendable with modifications to at least some of the other species of putative normative conflicts. But, even if this model has no application outside of theoretical normativity, the problem of comparing evidential and goodness-based reasons for belief is an important one in its own right.

4.3.1 Evidentialism, pragmatism, and the mixed view

This section introduces an example that will be referred to in the rest of the

chapter. After the example is introduced, there is an analysis of how three different normative theories of theoretical reason would begin evaluating the example. The first theory is epistemicism. Epistemicism is meant to be akin to evidentialism, and it is the view that there are only epistemic reasons for belief. The second view is pragmatism, the view that there are only goodness-based reasons for belief. The final view is the mixed view, which holds that there are both epistemic reasons and non-epistemic reasons for belief.

The example here concerns Jones. Jones comes from a family that left his hometown of Grodno, then in Russian-controlled Poland, when he was a boy. Now that Jones is an adult, the world has changed, and Grodno is no longer in Poland, but is part of Belarus. Jones, who does not follow world events, is unaware of this change of borders until he is given an atlas for his birthday. Shocked and saddened that his hometown is no longer in his native Poland, Jones falls into a deep depression. His only hope for recovery is not to have the belief that Grodno is in Belarus.

In evaluating this example, epistemicism would say that Jones ought to believe that Grodno is in Belarus. There is overwhelming evidence that Grodno is in Belarus. Jones's new atlas places Grodno in Belarus, other modern maps place it there, geography texts concerning the former USSR say that Grodno is in Belarus, and so on. On the other hand, there is very little evidence against the claim that Grodno is in Belarus. Weighing up the epistemic reasons, Jones ought to believe that Grodno is in Belarus.

By way of contrast, pragmatism would not be concerned with where the weight of the evidence was, except insofar as there is some benefit generally to believing what the

evidence suggests is true.[82] The pragmatist's primary concern would be to identify

whether or not there was more benefit in believing or not believing that Grodno is in

Belarus. In order to determine this, the pragmatist would have to weigh up the happiness

that Jones would gain from not believing that Grodno was in Belarus against the harm

done by having the false belief. That harm might include being unable to answer a

Belarussian geography question at a pub quiz, giving an incorrect answer to someone

who genuinely needed to know where Grodno was, and so on. Weighing up these facts,

Jones ought to believe or not believe according to the benefit.

The mixed view accepts that there are both epistemic and non-epistemic### in this

case goodness-based### reasons for belief. On the mixed view, the evidence counts in

favour of believing that Grodno is in Belarus. And, while some of the non-evidential

reasons would also weigh in favour of believing that Grodno is in Belarus, most weigh

against it.

Epistemicism and pragmatism share an important feature that is lacking in the

mixed view; they both operate on the assumption that there is just one source of

normativity for theoretical reasons. Although a complete theory of how to weigh benefits

is bound to be complex, as might be a complete theory epistemic considerations, there is

some intuitiveness to the idea that benefit can be weighed against benefit and epistemic

consideration against epistemic consideration. Evaluating what one ought to believe, all

things considered, within the mixed view does not intuitively seem as straightforward. It

is not immediately clear how epistemic considerations get weighed against benefits.

The rest of this chapter aims at illuminating how to relate evidential and non-

---

[82] I do not mean to suggest that this is a position with a significant philosophical
following.

evidential reasons for belief. The account is purely formal and does not offer to resolve

substantive questions about how much of one is enough to outweigh or defeat the other.

In illuminating how evidential and non-evidential reasons for belief relate to each other,

the rest of this chapter is also setting out a template for how different sources of

normativity, and possibly different domains of reason, can be compared to resolve

possible normative conflicts between them.

4.3.2 Ought to believe and reason to believe

Having now discussed an example in which there is an apparent normative

conflict between epistemic belief oughts and goodness-based (non-epistemic) belief

oughts, it is important to understand a feature of the way reasons and oughts relate. We

can begin by distinguishing between oughts and reasons. It may be tempting to take a

reason as a *pro tanto* ought. A *pro tanto* ought is an ought that holds, as long as there are

no countervailing oughts. If there are countervailing *pro tanto* oughts, then the first *pro*

tanto ought must be weighed up against them. A *pro tanto* ought just says that in the

absence of other *pro tanto* oughts, one ought to do or believe, and so on, what the *pro*

tanto ought says to do, believe, and so on. Taking reasons to be *pro tanto* oughts, if you

have a reason to believe *p* and no reason not to believe *p* and no reason to believe

something that implies ~*p*, then you ought to believe *p*. When other reasons are present,

perhaps some that favour believing ~*p*, then one is in the situation of its being the case

that one *pro tanto* ought to believe *p* and *pro tanto* ought to believe ~*p*. To find whether

one ought to believe *p* or ~*p* all things considered, one must weigh up the reasons for

believing *p* or ~*p* (or for remaining agnostic).

John Skorupski argues that there are two different reason operators.[83] One is the reason operator and one is the sufficient reason operator. On Skorupski's view, while there may be a reason for me to believe that it will rain tonight, it may not be a sufficient reason for me to believe that it will rain tonight. There has to be enough reason to believe something for there to be a sufficient reason to believe it. On this view, reasons cannot be regarded as *pro tanto* oughts. That is because there being a reason for belief, even when there are no reasons for not believing or for believing not, would not automatically make it the case that there is sufficient reason to believe *p*. If we use the weighing metaphor, we can say that the balance scale is preloaded such that a reason has to be of a certain weight to tip the scale towards the ought, even when there is nothing on the other side.

The use of weighing in determining what one ought to do, believe, feel, and so on does not apply in all cases. While we most commonly think of reasons as being weighed, there are cases in which reasons do not always combine by weighing. In certain circumstances, reasons that may initially have been candidates for weighing become irrelevant. In the case of epistemic and non-epistemic reasons for belief, when the non-epistemic reasons become strong enough, the epistemic reasons cease to play any role at all in determining what one ought to believe.

Returning to the Jones case, all the epistemic reasons available to him would require Jones to believe that Grodno is in Belarus. The non-epistemic reasons tell us that Jones ought to believe that Grodno is not in Belarus. If weighing is the correct way of resolving the putative normative conflict here, the problem would be set up like this.

Assume under some set of conditions *C*, the epistemic reasons outweigh the non-epistemic ones. So all things considered, Jones ought to believe that Grodno is in Belarus. If epistemic and non-epistemic reasons can be weighed against each other, the weight might fall to what Jones ought to believe according to the epistemic reasons. He has compelling epistemic reasons that come from maps, atlases, and geography tracts to tell him that Grodno is in Belarus. And, while there is some non-epistemic, goodness-based reason for Jones to believe that Grodno is not in Belarus, he will only suffer mildly for believing that Grodno is in Belarus.

Now consider another set of circumstances, *D,* in which the non-epistemic reasons are stronger than they were in *C*. Here, believing that Grodno is in Belarus is so deeply upsetting to Jones that he cannot reasonably manage in life. As long as Jones believes that Grodno is in Belarus, he will be unemployed, depressed, and unable to care for his family. One the other hand, if he does not believe that Grodno is in Belarus, he will become happy again and flourish. For sake of argument, assume that in *D*, the non-epistemic reasons outweigh the epistemic ones, but just barely. So, if some change in circumstances occurred that generated a new epistemic reason for Jones to believe that Grodno was in Belarus### perhaps Jones's being told by the world's leading authority on Eastern Europe that Grodno was in Belarus### then the balance could easily tip back towards its being the case that Jones ought, all things considered, to believe that Grodno is in Belarus. Some new benefit not to having the belief that Grodno is in Belarus might then come up, perhaps Jones will get a prize from his employer if he does not believe that Grodno is in Belarus, and this new benefit would add enough to the weight of the non-epistemic reasons to tip the all-things-considered ought back to what the non-epistemic

---

[83] Skorupski (forthcoming).

reasons weigh in favour of.

This delicate tipping back and forth of the balance of reasons fails to capture the way in which epistemic reasons and non-epistemic reasons interact. Instead it seems as though once non-epistemic reasons are being considered, the epistemic ones are no longer in play, even if they become stronger.[84]

To analyse how epistemic and non-epistemic reasons for belief relate, it will be helpful to look at an example. Recall the example in section 3.4. A basketball player, Smith, was introduced. Smith was in the unfortunate position that he had 5% less chance of making his free-throws than what he believed his chances to be. When Smith became aware of this, if only epistemic reasons counted, it would be the case that he ought to believe his chances of making a free throw were 0%. This was unfortunate for Smith, who was at the free-throw line with the outcome of the game depending on his making both of his free-throws.

Suppose that in addition to being a basketball player, Smith is a gambler and a philanthropist. He has placed a large bet against long odds that his team will win. If Smith wins the bet, he will donate all the proceeds to charity. If Smith forms his belief about his chances of making the free-throws based on the epistemic reasons alone, he is certain to miss, with the result that several people will fail to have their suffering alleviated for the loss of his donation. If he is confident that he will make the free-throw, even though his confidence would be unwarranted by the evidence, then he stands a much better chance of making at least one of the two free-throws, putting his team in a

---

[84] There is an important exception to this claim. Having one's beliefs track the evidence may be a benefit. On this view, when the amount of evidence increases, the amount of disvalue to believing something contrary to the evidence also increases. This might lead to a change in what one ought to believe, all things considered because of a change in the evidence. But, the change occurs only because the

position to win the game. Because something so significant is at stake in this circumstance than just winning a basketball game, it seems that no collection of epistemic reasons would be relevant in the face the overwhelmingly strong non-epistemic ones. The epistemic reasons, if we feel that they are not the only thing that count, seem under these circumstances to lose their force altogether.

That the epistemic reasons cease to count once they are silenced becomes evident in a thought experiment. Imagine that shortly before Smith was to shoot his free-throws, a coach runs onto the court, showing him a practice chart that puts his free-throw shooting percentage at 55%. Shortly thereafter, another coach brings to Smith's attention that his percentage has been 55% not just this season, but for the last three. Yet another coach bring a computer analysis of all the games that Smith has played, and this analysis says that Smith has shot 55% from the free-throw line. All this new evidence does not seem to matter in the face of the reasons derived from the benefits of alleviating the suffering.

On the other hand, if the benefits of making the free-throws were to decrease drastically, so would the benefits of Smith's believing that he is certain to make the free-throws. In this case, the epistemic reasons would not only come back into play, but they would also be the only reasons that did any work at all. The epistemic reasons are the only ones that count when the non-epistemic reasons are below a certain weight, but they do not count at all once the weight of the non-epistemic reasons are sufficient.

Here is an additional consideration against the view that the epistemic and non-epistemic reasons can be compared by a simple weighing procedure. Epistemic reasons, have a fixed range of weights. Evidence that gives a probability of 1 to a proposition

_____

disvalue of believing something contrary to the evidence has increased.

164

provides the maximum amount of epistemic reason possible for believing something; evidence that gives a probability of 0 to a proposition provides the maximum amount of epistemic reason not to believe it (or to disbelieve it). That you are in Culver City means that you are certain to be in Los Angeles County. No additional evidence, or any other epistemic consideration for that matter, gives you additional reason to believe that you are in Los Angeles County. Discovering, for example, that you are also three miles East of Santa Monica, California, does not provide more epistemic reason for you to believe that you are in Los Angeles County, even though being three-miles East of Santa Monica is certainly a sufficient condition for being in Los Angeles County.

This feature of epistemic reasons for belief makes weighing them against non-epistemic reasons for belief seem a little odd. Whether or not you ought to believe something, *qua* evidential reasons for belief, is not a matter of what there is more or even most reason for you to believe, it is a matter of whether or not there is *sufficient reason* for you to believe something. By way of contrast, with goodness based reasons for belief, what you ought to believe is just a function of what there is most reason for you to believe. Even a minimal amount of goodness-based reason, in the absence of any goodness-based reasons against, entail that you ought, *qua* goodness based reasons, to believe what there is that minimal reason for.

In general, however much reason there is for you to believe something on goodness grounds, there could always be more reason. And that additional reason might matter at some point. Presently there is a good deal of reason for me to believe I ought to pay my bill at the cafe where I am typing my thesis. The woman at the till just added a new reason for me to believe that I ought to pay, namely that she will phone the police if

I do not pay post-haste. Initially, it was the case that I ought to believe that I ought to pay my bill. Now that there is the threat of legal trouble, there is even more reason for me to believe that I ought to pay my bill (not to mention more reason that I ought actually to pay it!). This additional reason might matter if some reasons came up for me to believe that I ought to leave without paying my bill. Perhaps a friend of mine will make a big donation to Oxfam if I leave without paying my bill and then send in the money the next day. Before the threat of police involvement, my friend's offer might have been enough to make it so that I ought to believe that I ought not to pay my bill before leaving. After the threat of legal involvement, his offer was insufficient reason for me to believe that I ought to leave without paying my bill.

This asymmetry between evidential and non-evidential reasons makes the question of just what is being weighed somewhat vexing. Is one weighing up epistemic reasons versus goodness based ones? It seems like the relevant question is not whether there is more epistemic reason than goodness based reason, but rather whether one ought to believe what the epistemic reasons require or what non-epistemic reasons require, when what one requires is not compatible with what the other requires. To put it another way, it does not look like a comparative matter of whether or not there is more epistemic or non-epistemic reason when the two conflict, it is whether one ought to believe what the evidence and other epistemic considerations suggest or believe what the non-epistemic considerations say to believe.

4.3.3 Comparing evidential and non-evidential reasons for belief

The section provides a model for comparing epistemic and non-epistemic reasons. The model will be given in terms of the example of Smith. The model compares epistemic and non-epistemic reasons for belief on the assumption that when non-epistemic reasons for belief are strong enough, epistemic reasons for belief are silent, and that otherwise, non-epistemic reasons for belief are silent.

In its formulation here, the model for comparing epistemic and non-epistemic reasons in the free-throw case is a function that takes two arguments, $\{x, y\}$, that are mapped to two values, 'Ought to believe Smith is certain to make both free-throw attempts' and 'Ought not to believe that Smith is certain to make both free-throw attempts'. Let x be the weight of the epistemic reasons for believing that Smith is certain to make the free-throw. Let y be the weight of the non-epistemic reasons for believing the Smith is certain to make the free-throw.

Let x and y both take positive and negative values, and zero. When the value for x is positive, then the epistemic reasons are for believing that Smith is certain to make the free-throw. When the value for x is negative, then the epistemic reasons are against believing that Smith is certain to make the free-throw. If the value for x is zero, then the epistemic reasons are neither for nor against believing that Smith is certain to make the free-throw.

When the value for y is positive, then the non-epistemic reasons are for believing that Smith is certain to make the free-throw. When the value for y is negative, then the non-epistemic reasons are against believing that Smith is certain to make the free-throw. If the value for y is zero, then the non-epistemic reasons are neither for nor against believing that Smith is certain to make the free-throw. Then there is a negative number

and $\underline{y}$[85] and a non-negative number $\bar{x}$ and $\bar{y}$[86] such that:

If $\underline{y} < y < \bar{y}$, then if $x > \bar{x}$, Smith ought to believe that Smith is certain to make the free-throw, and if $x \leq \bar{x}$, Smith ought not to believe that Smith is certain to make the free-throw.

If $y \leq \underline{y}$, then Smith ought not to believe that Smith is certain to make the free-throw.

If $y \geq \bar{y}$, then Smith ought to believe that Smith is certain to make the free-throw.[87]

Depending on what values are set for upper and lower-limits of y, this sort of defeasing function makes it the normal case that we only consider evidential reasons in determining what we ought to believe. However, if it turns out to be very good or very bad to believe a particular proposition### that is to say, if the value of y is above or below it upper or lower limit respectively### then epistemic reasons are no longer in play. The only reasons that count are the non-epistemic ones.

This defeasing function captures some important features of how, if one subscribes to the mixed view, epistemic reasons and non-epistemic reasons might interact. It explains the silence of both epistemic and non-epistemic reasons under certain circumstances; if it is necessary to consider one epistemic reasons, then it is not necessary

---

[85] y lower limit
[86] x and y upper limit
[87] I am grateful to John Broome for providing the correct formalisation of this function

to consider non-epistemic reasons, and *vice-versa*. It also explains the sense of tipping or weighing that is involved### some additional strength of non-epistemic reasons can be enough to tip the balance towards non-epistemic reasons all together, and some loss of non-epistemic weight can just as quickly cause the balance to tip back to consideration of epistemic reasons. The tipping and silencing prevent its being the case that once it is appropriate to do what the non-epistemic reasons tell us to do, the balance might shift back to epistemic reasons if new epistemic reasons come to light. That was a principal concern of the weighing account.

4.3.4 Fixed and unfixed ranges for goodness

In the defeasing version of the mixed view, what we have essentially is a set of normal conditions### a certain range of values for the weight of the non-epistemic reasons### under which only epistemic reasons count in determining what we ought to believe in a particular case. When the value for the weight of our non-epistemic reasons is equal to or outwith the limits of the range, then only the non-epistemic reasons count in determining what we ought to believe.

If we are to adopt the mixed view and the defeasing function, something needs to be said about how the range for *y* is fixed. Is there a constant range for *y* or does the range change on a case-by-case basis? If it changes on a case-by-case basis, why does it change?

Consider the Smith example compared against another one of the same form. Jane lies in the hospital with a putatively terminal illness. It is a very close balance

for me.

between her surviving and dying. Jane needs all the help she can get to survive the illness. One thing that Jane can do is believe that she is certain to survive the disease. Having a positive attitude will bolster her immune system, thus increasing her chances of survival. Like Smith, Jane's belief about an outcome she desires will affect the likelihood of the outcome's occurring.

Given that her life is at stake and assuming that there are no negative consequences to so doing, it seems that Jane ought to believe she is certain to beat her illness, even if the evidence tells her not to be so confident. In terms of the defeasibility function, y would be above the upper limit of the range### i.e. the non-epistemic reasons weigh too heavily in favour of believing that she is certain to survive for the epistemic reasons to matter, even when the upper and lower limits of y are far apart.

Now consider a modified version of the Smith example. Smith is not playing for the championship, but instead is shooting free-throws on a bet with a friend. Smith will win a free and tasty lunch if he makes both of the free-throws, but he will have to buy his friend lunch if he misses. Smith's free-throw shooting ability is the same as before, and he is in the same epistemic state. Ought Smith to believe that he is certain to make both free-throws, taking on board the non-epistemic reasons and not considering the epistemic ones?

If we accept the defeasing function and the mixed view and think that Smith ought to believe that he is certain to make both of the free-throws, then we are committed to one of two positions. Either we think that the normal range for y is fixed and very narrow (since that something will result in a free lunch is not a very strong reason for it), or we think that the range for y is variable.

The consequence of accepting that the range for y is fixed and narrow in all situations is that epistemic reasons are easily overridden in all circumstances by non-epistemic ones. One advantage of the defeasing view seemed to be that it respected the suzerainty of epistemic reasons as reasons for belief except in unusual circumstances. When the range of 'normal' values for y is too narrow, 'normal' begins to describe such a narrow range of circumstances as to make 'normal' rather rare.

That leaves us with two choices as regards the modified Smith example. We can either say that he ought not to be certain that he will succeed, as the epistemic reasons require, or that we should adopt a varying range model of the function. I have no real argument against the former point of view, save to say that it seems implausible to me. If we do not accept the former view, then we are forced to consider the variability hypothesis. On this view, the range of the limits of y will change according to the circumstance. It is not clear that there is some set, formal algorithm that one could apply to determine the correct range. Rather it seems a substantive matter, requiring a kind of case-by-case evaluation.

However, it is worth looking at what sort of considerations might come into play. I shall mention one that seems particularly interesting. It is the question of how local or global the epistemic consequences will be of adopting the false belief, against the weight of epistemic reasons . In the case of the bet with a friend while playing basketball, the broader epistemic consequences are likely to be limited. Put simply, one is highly focused on a particular activity, one is unlikely to be reflecting in the sort of way that might cause this false belief to infect one's broader belief structure. It is not wholly clear that, even if a little bit of reflection is involved, much damage would be done to Smith's

overall belief structure. As a further point, Smith's belief is likely not to be long lasting. After he makes both free-throws, having believed against the epistemic reasons that he would make both of them, he might explain himself like this. He had a feeling; the ball felt good; he knew the statistics, but somehow they did not seem to matter. He is not sure what came over him, but he just knew he was going to make it. He knows that next time he probably will not do so well. If we accept the variable range view, it seems at least a plausible hypothesis that the range widens and narrows at least partially in proportion to considerations of how severe the epistemic consequences are of not going with the weight of epistemic reason.

4.3.5 Differences and similarities between defeasing and weighing

It is tempting to read the weighing view and the defeasing view as a reflection of a deep difference in the logical structure of reasons. That reading may well be valid. It is consistent with the defeasing view that there really is an ought that is not subject to normal weighing operations. This, however, need not be the case.

The primary reason to accept that there may be a defeasing relation between epistemic and non-epistemic reasons for belief is that not all normative operators are subject to weighing. Certain oughts, notably some of the ones generated by rational requirements, do not appear to be the result of having most reason to do something. It seems to be the default position that one ought to do what rationality requires.[88] It is possible that, as I argue, there are extraordinary circumstances under which one ought not

---

[88] This does not preclude the possibility that rational requirements *can* be weighed against other things.

to be rational,[89] but the fact that they are the exception only underscores the degree to which its being the case that one ought to do what is rationally required is the rule. It is more than a matter of convenience that we do not feel the need to weigh up whether or not to be rational in general.

Nevertheless, we can still interpret defeasing as, strictly speaking, a weighing operation. If we do so, we do so like this. Let us return to the case of epistemic and non-epistemic reasons for belief. The defeasing function does not itself have any metaphysical commitments, in this case it merely presents a mapping from the arguments of goodness-based reasons and epistemic reasons to the values 'Smith ought to believe that he will make the free-throw' and 'Smith ought not to believe that he will make the free-throw'. Why that mapping obtains is a separate matter. One reason why it might obtain is that it reflects a particular metaphysical reality### there are these defeasible, non-weighing oughts (in this case the oughts that arise from the weight of epistemic reasons), and they are defeated under such-and-such circumstances. Or, it might reflect some facts about weighing.

In the weighing interpretation, we could read the defeasing function like this. There is R much reason to do what the epistemic reasons tell us to do. R sets the limits for the amount of goodness### you need to have $R + 1$ weight of countervailing reasons to do something on some other basis. Arbitrarily, say we have a reason with a weight of 5 to believe what the epistemic reasons tell us to believe. So, we would need a goodness reason with a weight of 6 to outweigh that. On this reading, for a range of weights of goodness reasons (those between 5 and -5), the only thing that determines what we should believe is the mass of epistemic reasons. Outside this range, only the goodness

---

[89] See section 3.2.1.

reasons do the determining, since they will outweigh the reasons we have to look at the evidence. On this reading, the defeasing function is still, strictly speaking, a weighing function. What we weigh are not epistemic reasons against goodness based ones, but rather reasons to follow only epistemic reasons against reasons to follow only goodness based reasons.

There is a roughly parallel issue in trying to decide whether, in accepting the defeasing function, we are committed to the mixed view. The defeasing function is perfectly compatible with the goodness view, although not, it seems, the epistemic view. Suppose that on the goodness view, truth is valuable to a certain degree. So there is a particular weight of reason to believe what the evidence or other epistemic considerations suggests. However, it is only so valuable, so when it is better to believe something else, then we ought to believe something else. The goodness view could neatly account for varying range version of the defeasing function. Depending on how valuable following the epistemic considerations was in some particular case (which, by the way, might well correspond to how much damage would be done to your epistemic network for ignoring the epistemic considerations in a certain instance), it would be natural to expect the limits of when you ought to believe what the epistemic considerations suggest and when you ought to believe what is best to vary accordingly.[90] Note that this position would be conducive to a normative monist reasons generalist, whereas the mixed position is only compatible with normative pluralism and reasons generalism.

Conclusion

---

[90] Or to be more precise, when it is best to believe what the epistemic reasons say to believe or whether it is best to believe something else will vary accordingly.

This chapter discusses some of the ways that normative theories might understand and possibly resolve the putative normative conflicts discussed in chapter 3. The first part of this chapter discussed possible normative architectures. Looking at these architectures provides a clearer picture of what resources are available to normative theorists whose theories conform to these architectures.

Following a discussion of the architectures, an analysis was given of how putative normative conflicts could be taken to be only apparent normative conflicts in one of the architectures, normative pluralism and reasons generalism. This architecture was explored because it is commonly held that there is a plurality of sources of normativity. The arguments in chapters 1 and 2 of this thesis provide support for reasons generalism by casting some doubt on the plausibility of reasons specialism through the concerns they raised about normative separatism.

The method discussed for resolving apparent normative conflicts in NPRG was the defeasing function. The defeasing function is an account of the way in which evidential and non-evidential reasons may be compared to arrive at an all-things-considered ought concerning what an agent ought to believe. A function of roughly this form may be suitable for accounting for the structure of a variety of possible normative conflicts in an NPRG framework.

Conclusion

In this thesis, I have argued that putative conflicts of normativity are an important issue for normative theorists. A normative conflict occurs when two or more of the requirements of normativity are not mutually satisfiable. I have discussed three types of these conflicts in particular.

In one case, normativity tells you that you ought to comply with what rationality requires of you. Call this ought a rational ought (O_r). In this case, normativity also tells you that you ought not to comply with what rationality requires of you, in this instance on moral grounds. Call this ought a moral ought (O_m). This putative normative conflict has the form:

C1. $O_r p$ and $O_m \sim p$.

In another case, there is a putative conflict between two different sets of considerations about what to believe. Suppose normativity tells you that you ought to believe what the weight of the epistemic reasons points to. Call this ought the epistemic ought (O_e). Suppose normativity also tells you that you ought to believe what the weight of the non-epistemic reasons points to. Call this ought the non-epistemic ought (O_n). When these two oughts require you to have and not to have a particular belief, the putative normative conflict has the form:

C2. $O_e B q$ and $O_n \sim B q$

In the final case, there is a putative conflict between what you ought to believe and what you ought to do. Suppose you ought to believe x. Call this the theoretical ought (O_t). Suppose also that you ought to do y. Call this the practical ought (O_p). Suppose, too, that it is not possible that you both believe x and that you do y. In this case, you cannot comply with both the practical ought and the theoretical ought. This putative conflict would have the form:

C3. ($O_t B x$ and $O_p D y$) and ~($B x$ and $D y$)

The notation in C1, C2, and C3 employs oughts with subscripts. One of the key issues to come to grips with when considering putative normative conflicts is whether or not there is an overarching ought that has no subscript. A normative separatist accepts that the oughts with subscripts in C1 and C3 are the basic oughts, that there is no overarching notion under which the oughts with subscripts are subsumed. In C2, the normative separatist, if he were not an epistemicist, might accept that there is a basic concept of a belief ought that subsumes the epistemic and non-epistemic belief oughts. Epistemicists, of course, will not accept that there are non-epistemic belief oughts.

It may be tempting to think that if a normative theory accepts that at least one pair of these oughts with subscripts cannot be subsumed under some broader concept of ought, then that theory takes normative conflicts to be real, as there are two oughts that cannot be compared and whose requirements cannot both be met. However, this need not be the case. Theories that hold that there are multiple oughts can interpret the situation in

two different ways, only one of which constitutes accepting that conflicts are real.

The first interpretation is consistent with the view that there are no normative conflicts. This interpretation holds that it does not make sense to talk about normativity generally, but instead to speak of specific kinds of normativity### moral normativity, prudential normativity, practical normativity, theoretical normativity, and so on. While on this interpretation it would not be false to say that you might not in some circumstance be able to meet both the requirements of theoretical normativity and the requirements of practical normativity, this would not be a normative conflict. The reason why this would not be a normative conflict is that there is no concept of normativity that includes both theoretical normativity and practical normativity.

The second interpretation holds that there is some broader notion of normativity that underpins theoretical normativity, practical normativity, and so on. But, there is no more basic ought than that of theoretical normativity, practical normativity, and so on. On such a view, the requirements of normativity might conflict, and thus normative conflicts are real, not merely apparent. Whether or not the possibility of normative conflicts is a worry for a theorist who takes this view depends on whether or not the theorist regards normative conflicts as being problematic or not. A theorist who does not want normative conflicts either must argue that the conflicts discussed in this thesis do not actually arise or she must argue that even without an overarching concept of ought, her theory can resolve such conflicts, making them only apparent conflicts.

For theorists who accept that there is a basic notion of ought that underpins the oughts with subscripts in C1, C2, and C3, other issues arise. If normative conflicts are not permitted in a certain theory, then it must be shown that either the putative conflicts

discussed in chapter 3 of this thesis do not arise, or that they can be resolved.

I have argued in this thesis that various putative normative conflicts do arise. In chapter 1, I argued extensively against evidentialism. Although I did not argue explicitly against the view that there are only epistemic reasons for belief, I believe that it would be similarly hard to motivate this view. As long as non-epistemic reasons for belief are allowed, there is always the potential for a conflict between what an agent ought to believe, according to his epistemic reasons, and what an agent ought to believe, according to his non-epistemic reasons. A theory that does not accept that there are normative conflicts that arise from different sorts of reasons for belief must show one of two things. It must show either that different sorts of reasons for belief never conflict, or it must deny that one of the conflicting sorts of reasons for belief are genuinely reasons for belief.

For inter-normative conflicts, there is no way to avoid the fact that the requirements of practical normativity may conflict with those of theoretical normativity. Sometimes doing what you ought to do makes it so that you cannot believe what you ought to believe, and *vice versa*. For theorists who think that there is not an overarching concept of normativity that encompasses practical and theoretical normativity, this situation may not prove alarming. Nonetheless, it seems to me that there is reason for concern, whatever one's broader normative view. When there are several putative normative requirements on an agent, it seems odd to think that there is no choice that the agent can make, no way that the agent can be, so as not to be guilty of failing to comply with a normative requirement. If there is no concept of normativity independent of the various concepts of normativity within a certain domain, then there is no guidance for those cases in which doing what one ought to do makes it so that one cannot believe what

one ought to believe, and *vice versa*.

I have proposed to account for one apparent normative conflict, the conflict between epistemic and non-epistemic normative requirements to believe, with something that I have called the defeasing function. The defeasing function is intended to capture the complicated way in which some types of conflicting reasons for belief may be compared with each other. The defeasing function provides a mechanism whereby one sort of reason for belief can be the only sort that counts up until the point at which another sort of reason for belief is sufficiently strong, at which point it is only the second sort of reason for belief that counts. A function of this sort may be suitable for a much broader variety of cases, but it is my suspicion that the correct types of comparisons among different sorts of reasons or different domains of normativity may vary a great deal from case to case.

One type of conflict for which a good method of resolution is required is that between the normative requirement that we comply with a rational requirement and the moral requirement, which can arise in some circumstances, that we not comply with some rational requirement. While I have argued that rational requirements are not themselves normative requirements, I have argued that they may give rise to normative requirements. An account of how to compare the competing normative requirements arising from rationality and normativity, respectively, may take a similar form to the account of how to compare epistemic and non-epistemic reasons for belief. What the two cases have in common is that under normal circumstances, one consideration is the only one that does any work, while in extraordinary circumstances, that same consideration is silent. In the case of belief, epistemic reasons are the only ones that count, unless the

non-epistemic reasons are very strong, in which case the epistemic ones are silent. With rationality, it seems plausible to think that the normative requirement to comply with a rational requirement is the only normative requirement that is relevant, unless there are very strong (perhaps moral) reasons not to comply with a requirement of rationality, in which case the normative requirement arising from rationality is silent.

Putative normative conflicts place great demands on theories of normativity and even on theories of rationality, as discussed *vis a vis* Parfit's theory of rationality. Although his theory is not sufficiently well spelled out to be sure, it appears to be possible that on Parfit's theory of rationality, normative separatism would imply rational separatism. I have shown that Parfit's theory is not compatible with rational separatism, so one down-stream consequence of accepting normative separatism is that one cannot also accept Parfit's theory of rationality.

Whether a theory of normativity accepts that putative normative conflicts are real or maintains that they are only apparent, putative normative conflicts place constraints on that theory. This thesis has provided an account of how putative normative conflicts can arise, what some of the ramifications of them are for theories of normativity, and how some theories of normativity might begin addressing them.

References and Select Bibliography

Adler, J. (2002), *Belief's Own Ethics* (Cambridge, Mass.: Bradford Books).

Blackburn, S. (1984), *Spreading the Word* (Oxford: Oxford University Press).

—— (1998), *Ruling Passions* (Oxford: Clarendon Press).

Bratman, M. (1987), *Intention, Plans, and Practical Reason* (Cambridge, Mass.: Harvard University Press).

Broome, J. (2000), 'Normative Requirements', in J. Dancy (ed.), *Normativity* (Oxford: Blackwell Publishers Ltd).

—— (2001*a),* 'Are Intentions Reasons? And How Should We Cope with Incommensurable Values?', in C. Morris and A. Ripstein (eds.), *Practical Rationality and Preference: Essays for David Gauthier* (Cambridge: Cambridge University Press).

—— (2001*b*), 'Normative Practical Reasoning', *Proceedings of the Aristotelian Society*, Supplementary Volume 75, 175-93.

—— (2002), 'Practical Reasoning', in A. Millar, *Reason and Nature: Essays in the Theory of Rationality* (Oxford: Oxford University Press).

—— (2004), 'Reasons', in R. Jay Wallace, P. Pettit, S. Scheffler, and M. Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Oxford: Oxford University Press).

Chang, R. (1997), 'Introduction', in Chang (ed.), *Incommensurability, Incomparability, and Practical Reason*. (Cambridge, Mass.: Harvard University Press).

Cherniak, C. (1992) *Minimal Rationality* (Cambridge, Mass.: MIT Press).

Dancy, J. (1977), 'The Logical Conscience‘, *Analysis*, 37, 81-84.

— (2000), *Practical Reality* (Oxford: Oxford University Press).

Dennett, D. (1998) 'Brain Wiring and Mind Reading', in *Brainstorms* (Cambridge, Mass.: MIT Press).

Ewing, A. C. (1959), *Second Thoughts in Moral Philosophy* (London: Routledge and Kegan Paul).

Gibbard, A. (1990), *Wise Choices, Apt Feelings* (Cambridge, Mass.: Harvard University Press).

Hájek, A. (2004), 'Waging War on Pascal's Wager', *Philosophical Review*, 113, 27-56.

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

Harman, G. (1999), *Reasoning Meaning and Mind* (Oxford: Oxford University Press).

Hieronymi, P. (unpublished typescript), 'Controlling Attitudes'.

Hookway, C. (2000), 'Epistemic Norms and Theoretical Deliberation', in J. Dancy (ed.), *Normativity* (Oxford: Blackwell Publishers Ltd).

Kelly, T. (2002), 'The Rationality of Belief and Some Other Propositional Attitudes', *Philosophical Studies*, 110, 163-196.

Loewer, B. (1994), 'The Value of Truth', in E. Villanova (ed.), *Information and Content* (Atascadero, Calif.: Ridgeview Publishing Co.).

Mackie, J. L. (1977), *Ethics: Inventing Right and Wrong* (London: Penguin Books).

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

Nozick, R. (1993), *The Nature of Rationality* (Princeton, NJ: Princeton University Press).

Parfit, D. (1984), *Reasons and Persons* (Oxford: Oxford University Press).

— (2001), 'Rationality and Reasons', in D. Egonsson, J. Josefsson, B. Petersson, and T. Rønnow-Rasmussen (eds.), *Exploring Practical Philosophy* (Aldershot: Ashgate).

Pojman, L. P. (1985), 'Believing and Willing', *Canadian Journal of Philosophy*, 15, 37-55.

Pollock, J. (1979), 'A Plethora of Epistemological Theories', in G. Pappas (ed.), *Justification and Knowledge* (Dordrecht: Reidel).

Rabinowicz, W. and Rønnow-Rasmussen, Toni (forthcoming), 'The Strike of the Demon', *Ethics.*

Railton, P. (1994), 'Truth, Reason, and the Regulation of Belief', *Philosophical Issues*, 5, 71-94.

⸺ (1997), 'On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action', in G. Cullity and B. Gaut (eds.) *Ethics and Practical Reason* (Oxford: Oxford University Press)

Raz, J. (2000), 'Explaining Normativity: On Rationality and the Justification of Reason', in J. Dancy (ed.), *Normativity* (Oxford: Blackwell Publishers Ltd).

Ross, W. D. (1930), *The Right and the Good* (Oxford: Clarendon Press).

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

Skorupski, J. (1999), *Ethical Explorations* (Oxford: Oxford University Press).

⸺ (forthcoming), *The Domain of Reasons*.

Smith, M. (2003), 'Neutral and Relative Value after Moore', *Ethics Centenary Symposium on G. E. Moore's* Principia Ethica, 113, no. 3, 576-598.

Stich, S. (1993), *The Fragmentation of Reason* (Cambridge, Mass.: MIT Press).

Velleman, D. (2000), 'On the Aim of Belief', in Velleman, *The Possibility of Practical Reason* (Oxford: Clarendon Press).

Wedgwood, R. (2002*a*), 'Conceptual Role Semantics for Moral Terms', *Philosophical Review,* 110, 1-30.

— (2002*b*), 'The Aim of Belief'', *Philosophical Perspectives* 16, 267-297.

Williams, B. A. O. (1973), 'Deciding to Believe', in Williams, *Problems of the Self* (Cambridge: Cambridge University Press).

— (1981), *Moral Luck* (Cambridge: Cambridge University Press).

— (1985), *Ethics and the Limits of Philosophy* (London: Fontana Paperbacks).