PRACTICAL REASONS

A thesis submitted in partial fulfilment of the requirements

for the Degree

of Doctor of Philosophy in Philosophy

in the University of Canterbury

by Carolyn E. Mason

University of Canterbury

2012

Contents

Acknowledgements

Abstract

Normal human limitations mean that when people decide how to act, they often have to base their decisions on flawed information or reasoning. Even when agents reason to the best of their ability, and form intentions consistent with that reasoning, they sometimes get things wrong. Dominant theories about reasons for action argue that all good, or 'normative', reasons for acting are objective normative reasons. But objective normative reasons for action are derived from facts about the world that ignore certain facts about human agents. On these accounts of reasons, real human agents can be unable to learn what they have normative reason to do. A common response to this problem is to say that in such situations people act in a praiseworthy way, but their actions are based on false beliefs, and false beliefs cannot be good reasons. I argue that when agents reason to the best of their ability and form intentions consistent with that reasoning, agents act appropriately in response to states of the world that are normative reasons for action. To support my claim, I develop an account of what I call 'justifiable reasons', normative reasons for action that human agents can always use as a basis for action, and the form of reason that underpins rationality.

I discuss the similarities and differences between my account of justifiable reasons and several approaches to reasons that resemble my account. I show that, in spite of objections, justifiable reasons are normative reasons, not motivating reasons. Accounts of subjective normative reasons are based on examples that look similar to mine. So, I explain why justifiable reasons are not subjective normative reasons. Some features of internal reasons also resemble features of justifiable reasons. But, I show that there is nothing about justifiable reasons that entails that they must be internal or external reasons. I take it that justifiable and objective normative reasons serve different purposes, so I explain these different purposes. Finally, I argue in support of my claim that to be rational, agents must act appropriately in response to justifiable reasons.

1. Introduction

Here is my thesis:

> If I can act, I can always act for a good reason.
>
> Something is not a reason for me to act if I cannot become aware of it.
>
> I am rational when I act appropriately in response to good reasons.
>
> Apart from trivial cases where my actions are unimportant, I deserve blame when I fail to determine what I have good reason to do.

These statements assume that what people have good, or 'normative', reason to do depends on what they would be justified in taking themselves to have reason to do. Although these are commonplace claims among non-philosophers (try them on your non-philosophically minded friends), they are rejected by many of those who research practical reason. My response to these rejections is to develop a coherent, useful account of what I call 'justifiable reasons for action', normative reasons that people would be justified in taking themselves to have reason to act on.

Real human agents have limited abilities to learn facts about the world and limited reasoning abilities. On my account of justifiable reasons, agents' reasoning abilities and circumstances affect which states of affairs are justifiable reasons for agents, and what agents have justifiable reason to do. In brief:

> An agent has justifiable reason to act if and only if, were the agent to consider the circumstances in a way that is possible and appropriate, she would hold that some state of affairs somewhat favours her acting in that way.

When I write of the agent that 'she would hold that some state of affairs somewhat favours her acting in that way', I mean only that she would act as though some state of affairs somewhat favours her acting in that way. Acting as though some state of affairs somewhat favours her acting in that way could involve anything from responding positively if someone asks her if she has such a reason, to acting on that reason. The normativity of justifiable reasons comes from justificatory ideals rather than ideals associated with values or consequences. Hence the importance of the requirement that the agent consider the circumstances in a way that is possible and appropriate. 'Possible and appropriate' carries a huge load within my theory, but I argue that we have a satisfactory commonsense understanding of what it means for someone to be justified in her conclusions about what she has reason to do. I claim that states of affairs, for example, that there are chairs in a room, favour agents acting in certain ways, and that this role is not usually played by mental states, for example, beliefs about there being chairs in a room. This matters, because the initial response to my claims about justifiable reasons is usually the response that I am not talking about good, that is 'normative', reasons at all and, hence, justifiable reasons must be what are called 'motivating reasons', which are often thought to be agents' beliefs, whether true or

false, about what they have reason to do. However, my claim that it is states of affairs that favour an agent acting as she has justifiable reason to act is modified by other aspects of this reason relation. I argue that states of affairs only serve as reasons when they are appropriately related to the potential consequences of acting in a certain way and the positive and negative values of acting in that way. The word 'value' needs to be read very broadly, so that it includes, as well as moral values, prudential values, aesthetic values and other forms of value, no matter how minor; this point needs to be kept in mind throughout the thesis.

Justifiable reasons seem to me to serve an important purpose because they make sense of the sentences with which I began this thesis, and because they explain what real human agents ought to take themselves to have some reason to do. However, I argue that justifiable reasons also play an important role in our concepts of praise and blame, self-regulation and rationality, a role that competing accounts of normative reasons cannot play. People are only praised and blamed for failing to act for good reasons when they were capable of learning what those good reasons were. Similarly, an agent who acts in a way that causes harm to her and others is not irrational if she acted as it was possible and appropriate to expect her to act given the circumstances. Finally, the reasons that real human agents ought to take account of for self-regulation only include the reasons that are practically accessible to them. Justifiable reasons serve all three of these purposes.

The connection my initial statements make between what people would be justified in taking themselves to have reason to do and something's being a normative reason for acting is commonly ignored or rejected. Bernard Williams ignores this possibility when he develops his example of a gin and tonic drinker (1981a). Assume that someone wants a gin and tonic. He is holding a glass of petrol, but he thinks the glass contains gin. Does he have good reason to mix the stuff in his glass with tonic and drink it? Williams argues that: 'If an agent's conclusion in favour of a certain action is essentially based on a false belief … the agent has no reason to do that action, though he thinks that he has' (2001, p. 91). Williams doesn't consider the possibility that the G&T drinker might be justified in taking the glass to contain gin. It is unclear whether Williams would think the agent has good reason to drink petrol and tonic if the agent were justified in taking the glass to contain gin and tonic. Michael Smith agrees with Williams' claim that the person who wants a G&T does not have a normative reason to drink P&T (Smith, 1994, p. 98). Jonathan Dancy makes a related claim. Dancy gives an example where he makes a decision which has bad consequences, then writes:

> [If] I can later explain the choice I made by pointing out that there were some crucial facts that I happened quite reasonably to have got wrong….There is a sense of 'justify' in which I can be said to have justified what I did. But … the balance of reasons was [not] in favour of the action. (Dancy, 2000, p. 7)

According to Dancy, in such a situation the agent could correctly take himself to have reason to act in some way, but not have good reason for acting in that way.

Unsurprisingly, given the rejections of links between justification and normative reasons for action, the connections my initial statements make between normative reasons, rationality

and praiseworthiness are also commonly rejected. For example, Derek Parfit claims that rationality and normative reasons are unrelated:

> '[If] I believe falsely that my hotel is on fire, it may be rational for 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.' (Parfit, 2001: 17)

So according to Parfit, even if he has the best possible understanding of his circumstances, if there is no fire, his decision to jump is rational, but he does not have good reason to jump. John Hawthorne and Jason Stanley distinguish between normative reasons and rationality in a similar way:

> Suppose I am driving to a restaurant, when I come upon a fork in the road. I think it is somewhat more likely that the restaurant is to the left than to the right. Given that these are my only options, and … I do not have the opportunity to make a phone call or check a map, it is practically rational for me to take the left fork. (Hawthorne & Stanley, 2008, p. 581)

Hawthorne and Stanley continue by claiming that although it is rational for the driver to act as he would be most justified in acting, the driver has no normative reason to act in that way:

> [If the] evidential probability that the restaurant is on the left is sufficiently high ..., then what it is rational to do may very well be to go left…. But, on our view, it is not proper to treat the proposition that the restaurant is on the left as a reason for going left. (Hawthorne & Stanley, 2008, p. 581)

So, what an agent has reason to do and what it is rational for her to do are again pulled apart. Similar claims are made by others working on practical reasons and rationality whose positions are discussed later in the thesis (Wallace, 2007; Williams, 1981a).

Those who reject my initial statements do so because they accept an account of objective normative reasons; some claim that such reasons are the only form of normative reason for action, others that they are the foundational form of normative reason. Broadly speaking, 'objective normative reasons' provide objectively good reasons for actions. They are 'objective' reasons because they stem from the way the world is, and not usually from agents' psychological states.[1] Accounts of objective normative reasons tend to be accounts of reasons that are so highly idealised that agents would need supernatural perspectives on the world to consistently determine what they have objective normative reason to do.

The level of idealisation required by objective normative reasons means that such reasons can be inaccessible to agents. This is a problem for those who argue that objective normative reasons are the only form of normative reason for action. To see the barriers to being able to act on objective normative reasons, consider Hawthorne and Stanley's example again (2008, p. 581). When driving to a restaurant, I am faced with a choice between turning left or right. I have no way of checking which road the restaurant is on, but some dim, inexpressible memory means I think it is more likely to be on the road to the left. Assume that there is no

---

[1] The 'not usually' restriction is explained in §3.3.

other way for me to locate the restaurant. What do I have objective normative reason to do? On many accounts of objective normative reasons, I have reason to go down the fork that leads to the restaurant. Similarly, John Skorupski writes, 'if a truck has broken down around the bend and you'll avoid hitting it only by braking right now, then there is a reason for you to break right now, even though you haven't seen the truck and don't know that there is reason to slow down' (2009, p. 121). Skorupski claims that this inability to access reasons for action is one of the things that distinguishes reasons for believing from reasons for acting (2009, pp. 120-121). I do not argue that accounts of reasons that are undetectable to normal human agents are mistaken. However, I do argue that it is a mistake to hold that objective normative reasons are the only form of normative reason.

R. Jay Wallace is not convinced that situations where agents are unable to know what they have reason to do are a problem. He thinks that we can almost always work out what we have objective normative reason to do (Wallace, 2005, pers. comm., March). If mundane actions such as 'take the usual route to work' are left out of consideration, I am not convinced that we can usually base our actions on objective normative reasons. One objective normative reason for me to come to university today is that there will be no earthquakes in Christchurch today. But that is not an objective normative reason on which I could base my decision to come to work. Accounts of objective normative reasons usually require knowledge of facts such as whether there will be an earthquake in Christchurch today, and I do not know that this fact obtains. Whichever of Wallace and I are correct, if idealised objective normative reasons are the only form of normative reasons, it cannot be true that 'If I can act, I can always act for a good reason.'

Supporters of objective normative reasons have failed to see the potential of an account of justifiable reasons. Objective normative reasons have the virtue of being based in the world. Objective normative reasons researchers tend to investigate the kind of idealisation that is required to determine whether some state of the world is a normative reason for action. Those who argue that objective normative reasons are the only form of genuinely normative reasons argue that false beliefs cannot be normative reasons. They haven't seen that it is possible to give an account of normative reasons, of the kind that I argue for, namely justifiable reasons, that also rests on states of the world, but which includes, as justifiable reasons, all and only those states of affairs that agents would be justified in taking to be normative reasons.. Some of those who argue that objective normative reasons are the only form of genuinely normative reasons find the idea that an agent can have most normative reason to act in two different ways incoherent. What an agent has most justifiable reason to do, and what the agent has most objective normative reason to do can conflict; Hawthorne and Stanley's driver might have most justifiable reason to go left and most objective normative reason to go right. But such concerns about conflicting normative reasons rest on a flawed assumption that the conflict affects what agents have reason to do. It doesn't. When objective normative reasons and justifiable reasons favour differing actions, real human agents can only base their actions on justifiable reasons.

Objective normative reasons and justifiable reasons serve different purposes. Objective normative reasons are revealed using one of several possible forms of agent-independent idealisation. Justifiable reasons are reasons for action that real human agents can use, and ought to be disposed to use, to guide their actions. Determining what an agent has justifiable reason to do requires idealisation, but a different form of idealisation. If a philosopher's purpose in developing an account of reasons is to establish something about the relationship between value and action, then objective normative reasons serve a useful function: for with objective normative reasons you can hold steady the state of the world and the relationship between precisely specified actions and their potential outcomes, and then consider how value relates to what an agent has reason to do. In contrast, if one's aim is to discover the normative reasons that real agents ought to use as a guide to action,, one should focus on the agents' justifiable reasons.

The disconnect between objective normative reasons and real agents leads some supporters of objective normative reasons to introduce a second form of normative reason, 'subjective normative reasons'. Accounts of subjective normative reasons are more varied than accounts of objective normative reasons. Sometimes the term is used to refer to reasons that seem not to be normative at all (Schroeder, 2008c, p. 14; Sie, Slors, & Brink, 2004, p. 4). When the term 'subjective normative reason' is used to refer to genuinely normative reasons, it is usually taken to refer to situations where agents are justified in believing that they have objective normative reasons for acting (Cullity & Gaut, 1997, pp. 1-2; Joyce, 2001, p. 53; Wallace, 2003a). This approach to normative reasons is problematic. First, it introduces an ontological difference between subjective normative reasons and objective normative reasons; the former are beliefs, the latter states of affairs or true propositions. It seems best to avoid this if at all possible. It seems preferable to have the reasons that agents would ideally act on be the same kind of thing as the reasons that agents are justified in taking themselves to have reason to act on. Second, in many situations agents cannot form justified beliefs about what they have objective normative reason to do. Hawthorne and Stanley's driver cannot be justified in forming a belief that he has objective normative reason to take the left fork in the road. The driver is aware that he does not know the correct route. It would be peculiar to hold that his gut feeling that he has reason to drive to the left would justify him believing that he has an objective normative reason to drive to the left. In such cases, accounts of subjective normative reasons leave us no better off than accounts of objective normative reasons.

Like those who claim that objective normative reasons are the only form of normative reason, those who develop accounts of subjective normative reasons fail to see the potential of an account of justifiable reasons. These researchers assume that objective normative reasons are foundational and try to base subjective normative reasons upon them. For example, Garrett Cullity and Berys Gaut, write 'in seeking an account of normative practical reasons, it is objective normative reasons that will be our primary concern: from this an account of subjective ones will follow' (1997: 2). But, one of the points I argue for here is that what agents would be justified in taking themselves to have reason to do cannot be explained in terms of objective normative reasons. Human agency is constrained; we can take account of

those constraints and still develop a suitably idealised theory of agency, normative reasons and action.

My research is on reasons for acting, 'practical reasons', not reasons for believing, 'theoretical reasons'. However, what an agent has practical reason to do will both affect and be affected by what the agent has theoretical reason to believe, so one cannot be discussed without making assumptions about the other. I claim that normative reasons, by which I mean the features of the environment that are picked out and called 'reasons' for acting, are states of affairs. However, agents' beliefs, or the beliefs that it would be appropriate for agents to form, affect what agents have justifiable reason to do. At times I talk about the beliefs that it would be appropriate for agents to form about what they have reason to do, or talk about the actions that it would be appropriate for agents to take to correct their beliefs, which will, in turn, affect what they have reason to do. I do not assume that an agent needs to have, or be able to have, beliefs about what she has reason to do for her to have a justifiable reason to act; an agent can have a justifiable reason to act, and act for that reason, even when the agent cannot know that she has reason to act, or cannot form a belief that she has reason to act. But, sometimes an agent ought to get her beliefs in order before she forms an intention to act in some way. Perhaps Agnes believes that flying is more dangerous than driving. If she has time available and access to the Internet, she probably has reason to check her beliefs before she forms an intention to drive for ten hours rather than fly for one hour. But, notice that checking the accuracy of her beliefs requires forming intentions to act. The unavoidable interactions between practical and theoretical reasons mean that I cannot develop a theory of practical reasons for real human agents without sometimes discussing what agents have reason to believe or making assumptions about theoretical reasons.

The relationship between reasons for belief and reasons for action is outside the scope of this thesis. It is, however, worth noting that work on reasons for belief tends to assume that reasons for believing are more like my justifiable reasons than like objective or subjective normative reasons. So, the implications of my account of justifiable reasons for the relationship between practical and theoretical reason are worth exploring.

Some of the philosophical works that I discuss here focus primarily on the so-called 'Humean theory of reasons'. The Humean theory of reasons is important within debates about practical reason, and it is entangled with the issues that concern me. Although the truth or falsity of the Humean theory of reasons does not concern me, I sometimes discuss work that is focussed on the Humean theory of reasons. For example, in §6.1.2, I discuss Mark Schroeder's arguments about reasons in *Slaves of the Passions*, and Schroeder's principal concern in this book is the Humean theory of reasons (2008c). And in Chapter 7, I discuss internalism about reasons, which also relates to the Humean theory of reasons. So, I sometimes need to take elements of arguments about the Humean theory of reasons and use them for my own purposes.

David Hume wrote: 'Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them' ([1739-1740] 1978, p. 415). On the simplest reading of this statement, it is taken to be a claim that the only thing that reasoning

can do is help agents achieve goals that they are already motivated to pursue. This suggests that an agent needs to desire goals in order to have reason to pursue those goals. ('Desire' needs to be read broadly here, to include hopes, wishes, and any other psychological states that motivate agents to act.) Consider the example of a terminally ill person, Agnes, who will live longer if she takes a certain medication. Assume that the only thing Agnes wants, or could be brought to want, is to die as quickly as possible. If Hume is right, then no matter what her family and doctors say to her, she has no reason to take the life-extending drug.

Not everyone accepts the Humean theory of reasons. Some people argue that what a person has reason to do depends on what it is in her interest to do as well as, or instead of, what she wants to do. On this way of thinking about what agents have reason to do, if Agnes wants to die quickly, but her psychologist knows that she would feel happier if she spent more time with her family before she dies, she may have reason to take the life-extending medicine even if she has no desire to take it. Other people argue that what a person has reason to do is affected by what has value as well as, or instead of, what is desired. If values underpin reasons and life has intrinsic value, then even if Agnes wants to die quickly, and her pain means that it is in her interest to die quickly, she has reason to take the life-extending medicine.

Arguments about the Humean theory of reasons often use the word 'reason' in a way that differs from my usual uses of the word. Thus far I have been using the word 'reason' to refer to states of affairs. The word 'reason', as in 'the reason for her action', can also be used to refer to some passion, desire, interest or value thought to give the agent reason to act in some particular way. So, we might say that Agnes's reason for not taking the medicine was her desire to die as quickly as possible. Or we might say that Agnes had reason to take the medicine because life has intrinsic value. As I explain in §3.2, I intend to avoid committing myself to any particular position about whether desires, interests, values, or some combination of the three, serve as reasons in this sense of the word. When I use the word 'reason' in the sense of the word picked out by 'the reason for her action', I use it to refer to a state of affairs that gives the agent reason to act in some particular way. So we might say that Agnes had reason to take, or not take, the medicine because her doctor said that it would extend her life. In this example, the doctor's advice is a state of affairs that serves as Agnes's reason. Whether it is reason not to take it because Agnes wants to die, or a reason to take it because life has intrinsic value, is not at issue. I return to this distinction between uses of the word 'reason' at various points throughout the thesis. However, my account of justifiable reasons does have some implications for Humean and anti-Humean theories of reasons, because consistency requires that justifiable reasons cannot be based on any values that agents would not be justified in taking to be valuable.

In summary, in this thesis I argue that the normative reasons that give real agents reasons to act are best explained in terms of my concept of justifiable reasons: that, on one sense of the word 'reason', justifiable reasons are states of affairs, and that justifiable reasons are normative reasons that serve as a guide to action when people engage in appropriate self-

regulation, when their actions are praiseworthy, and when they act rationally. An outline of the content of the chapters in this thesis is given below.

'Chapter 2: Reasons: Motivating, Normative, Pro tanto and Overall' My arguments about good reasons for action, whether they be justifiable reasons, subjective normative reasons, or objective normative reasons, normally take such reasons to be pro tanto normative reasons for action. So, before explaining the distinctions between justifiable, objective, and subjective normative reasons I explain what it means for something to be a pro tanto normative reason. The first section of Chapter 2, §2.1, explains the distinction and relationship between motivating and normative reasons. Understanding the distinction between motivating and normative reasons is particularly important for understanding objections to my account of justifiable reasons discussed later in the thesis. I argue that motivating and normative reasons are usually the same kind of thing, but that when we discuss motivating reasons we give a descriptive account of reasons, whereas when we discuss normative reasons we focus on the normativity of reasons. The next section, §2.2, introduces the concepts of prima facie, pro tanto and overall reasons. These concepts of reasons are in standard use within philosophy, but different philosophers use different terms to refer to them, and have different understandings of these concepts. After explaining the difference between overall, pro tanto and prima facie reasons, I briefly discuss the relationship between overall and pro tanto reasons and the 'ought' associated with reasons for action. I argue, first, that it is too hasty to conclude from it being the case that someone has overall reason to act in some way that she ought to act in that way, and, second, that agents ought to be disposed to take pro tanto reasons to favour them acting in certain ways, even when the agents have overall reasons to act in a different way. This conclusion influences my later account of justifiable reasons. Jonathan Dancy's work on reasons is frequently discussed in this thesis, so it is important to clarify the relationship between Dancy's reason terms and mine. I argue that Dancy's 'contributory reasons' are pro tanto reasons. I also use Dancy's claim that overall reasons are not 'count noun' reasons to introduce a common confusion about reasons and reason relations discussed in the next chapter.

'Chapter 3: Reasoning, Reason Relations, and Reasons' It is not possible to understand my concept of justifiable reasons, or the distinction between justifiable reasons and other reason concepts, without understanding the different uses of the word 'reason'. In this chapter, I clarify the most important uses of the word. We commonly speak of 'using reason', 'taking something to be a reason', or of 'having reason'. Although these three uses of 'reason' are related, the word refers to something different in each of these phrases. In the brief section on reasoning, §3.1, I discuss the relationship between intentional action and reasoning. This section argues that agents can have both motivating and normative reasons for acting without having engaged in reasoning that has led them to form beliefs about what they have reason to do. The ambiguity between something's being a reason and someone having a reason introduces confusion into some arguments about reasons. Reason relations are what we refer to when we speak of an agent having reason to act. In §3.2, I present my analysis of reason relations. I argue that it can be equally appropriate to refer to several different relata within the reason relation as 'reasons'. In this section I also argue that reasons are only reasons

because of their roles in reason relations. In §3.3, I explain which of the various possible relata that can be referred to as a 'reason' I focus on in this thesis. I argue that such reasons are normally states of affairs rather than beliefs or propositions. In taking this position, I agree with Jonathan Dancy. However, I show that motivating reasons can be taken to be states of affairs rather than beliefs, without the need to resort to Dancy's conclusion that reason explanations are non-factive. This matters because the problems that lead Dancy to introduce non-factive explanations of motivating reasons could also be used to raise objections to my account of justifiable reasons.

'**Chapter 4: Justifiable Reasons'** Chapter 4 introduces and explains the concept of normative reasons and normative reason relations that is the prime focus of this thesis. Justifiable reasons are normative reasons that it is possible and appropriate for agents to take themselves to have reason to act on. Approaches to reasons that resemble justifiable reason relations are usually said not to be normative. So, I show that justifiable reasons are normative and that the notion is coherent, and important within both real life and philosophy. Unlike other forms of normative reasons, justifiable reasons play a vital role in self-regulation, praise, blame and rationality, and in §4.1 I briefly explain why this is the case. The general requirements for justifiable reasons and reason relations are outlined in §4.2. My concept of justifiable reasons relies on the notion of possible and appropriate practical reasoning, so in §4.3 I explain what I mean by 'possible and appropriate practical reasoning', and explain the consequences of using this form of reasoning as a basis for normative reasons. In §4.3 I also argue that justifiable reasons are, ontologically, the same kind of thing as other forms of reasons; that is, one and the same reason can be a motivating reason, justifiable reason and objective normative reason for carrying out an action.

'**Chapter 5: Objections to Justifiable Reasons'** In this chapter I examine four objections to my account of justifiable reasons. First, in §5.1, I consider an argument that is thought to show that first-personal accounts of reasons cannot provide adequate accounts of practical reasons. According to this argument, no first-personal account of reasons can succeed, because agents sometimes act for good reason without having any conscious awareness of their reasons for acting. In response, I show that justifiable reasons are not first-personal reasons. Next, in §5.2, I consider claims that justifiable reasons and reason relations cannot be normative, because they cannot be based on appropriate truths or perfect reasoning. In response, I argue that justifiable reasons are as much states of affairs as objective normative reasons, and show that justifiable reason relations need not be based on facts and perfect reasoning for them to be normative. Finally, in §5.3, I consider two arguments based on Jonathan Dancy's work. First, I show that although Dancy gives an account of reasons that overlaps with my account of justifiable reasons, Dancy's objective normative reasons differ significantly from justifiable reasons. Second, I consider Dancy's argument that objective normative reasons and justifiable reasons cannot both be normative reasons, because if they were, statements such as 'he has normative reason to do something he has no normative reason to do' could be true. This strikes some people as nonsensical, yet I argue that such statements can be true.

'Chapter 6: Subjective Normative Reasons' Like me, those who develop accounts of subjective normative reasons develop their theories partly because of the inability of objective normative reasons to connect with real agents' decision making. This resemblance to my theory is reinforced by the subjective normative reasons theorists' use of examples that resemble mine. So, in this chapter I clarify the distinction between subjective normative reasons and justifiable reasons and argue that justifiable reasons achieve what the account of subjective normative reasons described here cannot achieve. The accounts of subjective normative reasons described in §6.1 seem to be hybrids of objective normative reasons and motivating reasons. The creation of these accounts of reasons seems partly motivated by a desire to explain situations where real agents' reasons do not correspond with ideal agents' reasons. But, unlike my concept of justifiable reasons, the concept of subjective normative reasons developed cannot serve the required purpose. The accounts of subjective normative reasons described in §6.2 bear a greater resemblance to justifiable reasons. But, on these accounts, such reasons amount to beliefs about objective reasons. In §6.3, I argue that this is a mistake. Finally, in §6.4, I describe an account of subjective reasons that avoids most of the problems afflicting the other accounts of subjective normative reasons in this chapter. However, this last theory provides an account of internal reasons (reasons limited by what agents could potentially be motivated to do), and the distinction between internal and external reasons raises a different set of issues from the accounts of objective, subjective and justifiable reasons discussed so far in the thesis.

'Chapter 7: Internalism and Justifiable Reasons' My account of justifiable reasons relies on arguments and motivations that resemble some of those used to support accounts of internal reasons. But, in §7.1 I show that the internal-external reasons distinction was created in response to a different set of issues than those addressed by the justifiable, subjective, and objective normative reasons distinctions. I show that my arguments in support of justifiable reasons neither presuppose nor support internalism about reasons, and that justifiable reasons are not internal reasons. Arguments given by two proponents of internal reasons, Bernard Williams and Michael Smith, seem to suggest that my account of justifiable reasons cannot be an account of normative reasons. In §7.2, I argue that Williams' and Smith's arguments do not show that justifiable reasons are not normative. Finally, in §7.3, I show that even though some of Williams' arguments about the nature of normative reasons could be taken to be criticisms of my position, his approach to reasons is generally compatible with my notion of justifiable reasons.

'Chapter 8: Objective Normative Reasons' Throughout this thesis, I state that justifiable reasons do not replace or compete with objective normative reasons, and that these two forms of reasons serve different functions. Chapter 8 explains the function of objective normative reasons, and it provides evidence for my claim that the accessibility of objective normative reasons to real human agents is typically irrelevant to theories about objective normative reasons. This second aim is met in §8.1, where I explain the arguments of a philosopher who explicitly takes the accessibility of objective normative reasons to be irrelevant to the existence of such reasons, and those of a philosopher with a more standard approach who acknowledges that objective normative reasons will often be inaccessible. In §8.2, I argue

that the various conceptions of objective normative reasons that philosophers have developed all aim, at least in part, to provide an account of reasons whose normativity derives from the relationship between reasons and values, where the word 'value' is, as always, read very broadly. This makes the concept of objective normative reasons a useful philosophical tool, and a tool that can achieve things that justifiable reasons cannot achieve. In §8.3, I argue that objective normative reasons are also a useful practical tool. Although justifiable reasons cannot be reduced to objective normative reasons, if an agent is justified in believing that she has an objective normative reason to act in some way, the agent has a justifiable reason to act in that way. This means that anyone who develops a correct account of objective normative reasons can improve our understanding of when agents have justifiable reasons to act.

'Chapter 9: Rationality and Reasons' In this chapter, I argue that in one important sense of the term, agents act rationally if and only if they act appropriately in response to justifiable reasons. In Chapter 4, I explained that people who argue that only objective normative reasons are normative reasons claim that theories resembling my account of justifiable reasons provide accounts of rationality or praiseworthiness not accounts of normative reasons. Up until this point in the thesis, my principal aim has been to support my claim that justifiable reasons provide an account of the normative reasons that real agents ought to use as a guide for action. normative reason. In this chapter, I argue in support of my account of rationality. After broadly explaining the form of rationality I am concerned with in §9.1 and §9.2, I consider the relationship between my account of rationality and accounts of rationality that resemble mine but have been shown to be implausible. In §9.3 I explain why the claim that rationality involves appropriately responding to reasons is commonly rejected, and why objections to this account of rationality do not undermine my claim that rationality requires appropriately responding to justifiable reasons. In §9.4 I consider claims that rationality is best thought of as requiring an appropriate response to beliefs, or justified beliefs, about reasons. The claim that rationality requires an appropriate response to beliefs about reasons is commonly rejected for falling foul of the so-called 'bootstrapping objection'. In §9.4.1, I show that this objection cannot be levelled against my position on the relationship between justifiable reasons and rationality. The claim that rationality requires an appropriate response to justified beliefs about reasons does not succumb to the bootstrapping objection, but could be criticised for conflicting with other requirements of rationality. This matters because I agree that agents act rationally when they respond appropriately to their justified beliefs about what they have reason to do. In §9.4.2, I respond to this objection. A more significant problem for my account of justifiable reasons arises from arguments that conclude that rationality is a relationship between cognitive attitudes such as beliefs, desires and intentions, rather than a relationship between the world and actions. I consider the implications of these arguments in §9.5, where I discuss the normative reasons that agents can have when deceived by evil demons.

 'Chapter 10: Summing Up and Moving Forward' This chapter summarises the main conclusions drawn in this thesis and briefly explains some of the implications of my arguments.

2. Reasons: Motivating, Normative, Pro Tanto and Overall

2.0 Introduction

Philosophy of action, like other areas of philosophy, has its own terminology and key conceptual distinctions. Philosophers discuss reasons for actions in ways that are unusual in everyday life, and invent or appropriate words to serve their purposes. The situation is complicated by philosophers' disagreements about the referents of terms and differing analyses of concepts. Given that I add my own peculiar word usage and concepts to the ones currently in use, it is important to clarify my use of various terms.

One of the most important conceptual distinctions within philosophy of action is the distinction between the reasons that actually guide an agent's actions, commonly called 'motivating reasons', and the reasons on which agents' actions should be based, 'normative reasons'. In brief, motivating reasons explain why an agent acted, or acts, in a certain way, while normative reasons are good reasons for acting. Imagine that you are travelling on a bus and the man next to you collapses. If you decide not to help him because you have hay fever and you think that any form of ailment is a reason to stay away from people who may be ill, then the reason that explains your action, your motivating reason, is that you have hay fever. However, this is not a good reason for you to decide not to help him; that is, your hay fever was not a normative reason for you not to help him. Reasons for action can be both motivating and normative. If you decide to help him because he has collapsed, and his collapse was a good reason for you to help him, then his collapse was both your motivating reason for acting and a normative reason for you to act.

I claim that justifiable reasons are normative reasons. A common objection to accounts of reasons that resemble my account of justifiable reasons is that they are accounts of motivating, rather than normative, reasons. In brief, someone might argue that because justifiable reasons are always accessible to agents, justifiable reasons cannot be normative reasons and must, therefore, be motivating reasons. In later chapters, particularly Chapters 4 and 5, I show that justifiable reasons differ significantly from motivating reasons, and in Chapter 6, I show this objection sometimes applies against accounts of so-called 'subjective normative reasons'. Understanding this objection, and my response to this objection, requires an understanding of the motivating–normative reasons distinction. This distinction is explained in more detail in §2.1, developed in later sections of this chapter, and revisited in later chapters.

Sometimes agents clearly have most reason to act in one particular way. If you are travelling on a bus and it gets to your stop without incident, you normally have most reason to get off the bus. However, even when an agent seems to have overwhelming reason to act in one way, she may have some reason to carry out other incompatible actions. Perhaps, even though the bus has arrived at your stop, you take the interesting conversation you are having with a fellow traveller to give you some reason to stay on the bus. A range of possible kinds of

reasons for acting have been suggested to account for the strength with which a reason or group of reasons favour an action. In §2.2, I explain the concepts referred to by the terms 'overall', 'pro tanto', 'prima facie' and 'contributory' reasons. Not all philosophers believe that all of these terms refer to real phenomena, but arguments about the existence of these types of reasons are beyond the scope of this thesis. The aim of §2.2 is only to ensure that the reader understands the concepts of reasons that are at issue in the remainder of the thesis.

2.1 Motivating and normative reasons

Return to the scenario where you are travelling on a bus. The man next to you collapses, and you decide not to help him. Consider two questions that could be asked about your reasons for acting: 'Why did you act in that way?' and 'What favours you acting in that way?'. 'Why did you act in that way?' asks for the reasons that explain your action, one form of which is called 'motivating reasons'. The reasons why someone acted as she did can be evaluated; so, motivating reasons can be normatively good or bad reasons for acting. 'What favours you acting in that way?' asks about the reasons that, in some sense or other, indicate that there is something to be said in support of someone in your situation acting in some way. These are usually called 'normative reasons'. Almost every form of normative reason that I refer to in this thesis could also be a motivating reason.

The question 'Why did you act in that way?' asks for an explanation of what lay behind your action. If you thought about your runny nose and based your action on that state of affairs, then the fact that your nose was runny was among the reasons that motivated your decision not to help the collapsed person. There were probably other motivating reasons in play at the same time. Hopefully, you also considered the state of the collapsed man and gave his collapse some weight when deciding how to act. In this way, motivating reasons play a role in the mental processing agents engage in when they decide what they have reason to do.[2] The reasons mentioned in response to questions about why an action was carried out are sometimes called 'explanatory reasons' as they explain why something happened. However, many writers claim that explanatory reasons include both reasons that have figured in an individual's reasoning and reasons that have played some other causal role that supported the agent's action (Audi, 2001, p. 119; Crisp, 2006, p. 108; Dancy, 2000, p. 7; Darwall, 2006a, p. 285; Macklin, 1968; Schroeder, 2008c, p. 11). Assume that when the man on the bus collapsed, you pulled out a tissue and blew your nose to signal to others that you were focussed on your runny nose rather than on the person who collapsed. One explanation for your pulling out a tissue might be that you were given a packet of tissues the day before. If so, the fact that you were given a packet of tissues is among the reasons that explain why you acted as you did. So, the fact that you were given a packet of tissues might be taken to be a so-called 'explanatory reason' for your action. The use of the term 'explanatory reason' to refer to these background conditions means that it is unwise to use the term 'explanatory reason' as a synonym for 'motivating reason'. It is unlikely that your receipt of a free packet of tissues played any role, conscious or unconscious, in the mental processing that lay behind

---

[2] Donald Davidson refers to motivating reasons as 'reasons that rationalize' actions (1963, p. 685).

your decision to act. In other words, the fact that you were given a free packet of tissues was an explanatory reason, but not a motivating reason.

It might be thought that a reason is only a motivating reason when it explains an action that has been performed. Like many other philosophers writing about reasons I see no need to accept this restriction (Dancy, 2000, p. 4; 2004, p. 21; Fantl & McGrath, 2009, p. 137; Smith, 1994, p. 96).[3] So-called 'motivating reasons' are reasons that dispose an agent to act rather than reasons that motivated a performed action. As Dancy puts it: '"Motivate" here is not the success-term which means "get to act" but the attempt-term which means "incline to act"' (2004, p. 21). There are two ways in which a motivating reason may fail to result in the performance of an action. First, an agent may deliberate about her circumstances and form an intention to act in some way, yet be unable to act as she intended due to obstacles that may be mental or physical and internal or external. You might, for example, take the collapse of the man next to you to be a reason to help him, but be unable to do so because you fall and hurt yourself, can't work out what kind of help he needs, or someone else helps him first. When something prevents the agent from acting as she intended, the agent's disposition to act is sufficient for the agent to be described as having had a motivating reason for acting. Second, I assume that there can be motivating reasons that are outweighed by other reasons, that is, reasons that are an important part of an agent's practical reasoning process, but which the agent never decides to act on. This idea that there are so-called 'pro tanto' motivating reasons is discussed further in §2.2. It is odd to call a reason that never motivated an action a 'motivating reason', but substituting a new term, such as 'motivational reason' seems clumsy and unnecessary. So, like Smith, Dancy, and Schroeder, I use the existing terminology.

'What favours you acting in that way?' asks a different question from 'Why did you act in that way?', and answering it requires reference to normative reasons. If you moved away from the person who collapsed next to you because your runny nose led you to assume that you had a cold, your reasoning and the reasons that motivated you to act as you did could have been mistaken. Your beliefs about the state of the world could have been false; you might have had hay fever rather than a cold and so had no reason to move away. The conclusions you drew about the consequences of acting in various ways could also have been mistaken. Assume that the person faked a collapse in a misguided attempt to get attention, and that helping him would encourage his bad behaviour. In such a situation, you might have no reason to help him. 'What favours you acting in that way?' asks why you have good reason to act some way. To answer this question, you need to think correctly about the features of the world that it was appropriate for you to take to be reasons for acting. These features of the world are called 'normative' reasons for action. In this example, they might be such states of affairs as 'I have hay fever' and 'he is not ill'. Normative reasons are part of the explanation of why it would be good for agents to act in certain ways.

---

[3] Schroeder takes a slightly different approach, arguing that motivating reasons are a subset of what he takes to be a larger set of reasons that corresponds to what I am here calling 'motivating reasons' (2008a, p. 12). This different use of terms has no effect on the discussion of Schroeder's work later in this thesis.

There is a sense in which motivating reasons are always normative reasons. I take it that motivating reasons are descriptive reasons; they describe motivationally important factors that contribute to an explanation of why an agent acted as she did. The same factors that contribute to an explanation of why an agent acted as she did could also be discussed in terms of their normativity. 'Normative reason' is usually taken to mean normatively good reason, but normatively bad reasons are also normative reasons. So, reasons can be motivating reasons without being normative reasons in the usual 'good reasons' sense of the term. This way of thinking about the distinction between motivating and normative reasons mirrors the standard distinction in philosophy between the descriptive and the normative. 'Descriptive' describes what is the case. 'Normative' describes what ought to be the case. A philosopher of science can take an example of a scientist at work and describe what the scientist did, or discuss the same case in terms of what the scientist ought to have done. Similarly, a philosopher of practical reason can take an example of an agent acting and use motivating reasons to describe one aspect of what explains her action, or normative reasons to explain what she ought to have taken herself to have some reason to do.[4] The motivating reasons that explain people's actions are a product of who they are, where they are, what matters to them, and of situational factors, like their moods and what they are focussed on at a point in time. What someone takes her motivating reasons to be affects how we should treat her. Independently of whether you helped the person who collapsed, the degree to which you were motivated to help says important things about you, and is normatively evaluable. In cases where an agent acts for a reason that is not a normatively good reason, I assume that the agent's reason for acting still has some normative quality, for example it might be normatively bad or normatively neutral.

Although a reason cannot be a motivating reason without being normatively evaluable, a reason can be a normative reason without being a motivating reason; that is, reasons can be normatively evaluable even though they are never causally involved in any action or in practical reasoning. Although reasons can be normative without motivating, some philosophers claim that normative reasons must be the kind of thing that could potentially explain why someone acted. In other words, normative reasons must be the kind of thing that could serve as a motivating reason. Philosophers with a variety of views about the nature of reasons for action have argued in support of this claim; it lies behind Jonathan Dancy's arguments that motivating reasons and normative reasons must be the same kind of thing, and Bernard Williams' arguments that for a normative reason to be a reason for an agent to act in some way, it must be appropriately connected to the agent's motivational set (Dancy, 2000, p. 101; Williams, 1981a, 1995a). In general I agree with Dancy and Williams that normative reasons must potentially be able to explain actions. However, some philosophers develop accounts of normative reasons that are inaccessible to real human agents. In Chapter 8, I describe one such account, and argue that such highly idealised accounts of normative reasons are useful even though they describe normative reasons that may not be able to motivate.

---

[4] Stephen Finlay suggests that normative reasons are 'reasons that support a claim that we *ought* to act' (2006, p. 1, italics in original, see also pp. 3, 6). This seems too strong, or so I argue in §2.2.

2.2 Overall, pro tanto, prima facie and contributory reasons

Whether they are motivating or normative, reasons for action rarely come on their own, and rarely favour acting in only one way. This means that we can distinguish between what an agent has most reason to do and between various other reasons that the agent has for acting. This section begins by explaining the terms used to refer to what an agent has most reason to do. I then briefly argue that an agent can have most reason to act in some way without it being the case that the agent ought to act in that way. Next, I discuss normative reasons that may not give an agent most reason to act. I argue that agents ought to be disposed to take such normative reasons to give them reasons to act, even when they have most reason to act in a different way. Jonathan Dancy has developed a theory of reasons that seems to resemble mine, so to clarify the relationship between Dancy's theory and mine I show that Dancy's contributory reasons correspond to the kind of reasons I discuss in most of the rest of the thesis. Dancy's work is also used to introduce a difference between reasons and reason relations that is discussed in the next chapter. Finally, I return to the discussion of the relationship between motivating and normative reasons I began in the last section.

Sometimes the state of the world means that an agent clearly has an overwhelming reason to act in one particular way. If you are in an empty shed and it catches fire, you probably have reason to do only one thing. Leave! A variety of terms have been used to refer to the reasons associated with what an agent has most reason to do. For example, Alan Goldman describes agents who have most reason to act in some way as having 'overriding reasons', while Stephen Darwall refers to 'sufficient' or 'all things considered' reasons (Darwall, 2006b, p. 28; Goldman, 2009, p. 10). When reasons give agents most reason to act in a certain way, I refer to these as the agents' 'overall reasons' for acting. When agents act for particular reasons, these are the agents' overall motivating reasons for acting. Agents' overall normative reasons for acting are the reasons that it would be best for the agents to act in response to after taking account of all aspects of their situations.

What a person has overall normative reason to do might be thought to equate to what that person ought to do, whether that 'ought' is a prudential or moral ought. However, this is not obviously true. An agent might have an overall prudential reason to take the quickest route to the shops, but it seems excessively demanding to say that she ought to take the quickest route to the shops. Similarly, it may be excessively demanding to insist that an agent with an overall moral reason to act some way ought to act in that way. Perhaps the agent has overall reason to act in a way that involves sacrificing her life, and such a sacrifice is considered supererogatory, that is, morally praiseworthy, but not obligatory. The question of whether agents ought rationally to act as they have overall reason to act is less straightforward. However, in general I assume that an agent can have overall reason to act in some way

without it being the case that the agent ought to act in that way.[5] John Broome calls reasons that agents ought to act on 'perfect reasons' (2006, pp. 34-35). According to Broome, to say that an agent ought to act in a certain way is equivalent to saying that the agent has a perfect reason to act in that way (2006, p. 35). Hence, I assume that Broome's 'perfect reasons' are a subset of overall reasons.[6]

Practical reasoning seems ultimately aimed at determining the action, or actions, that an agent has overall reason to perform. The most obvious candidates are that practical reasoning should lead to morally right action, whatever that is taken to be, or that it should lead to prudentially right action, again, whatever that is taken to be. However, there might also be aesthetically right actions, scientifically right actions, and so on. Evaluating all the different calls upon an agent to act, and the interactions between those calls for action, would be difficult. No decisive argument has been given to show that there is one ultimate goal of practical reasoning, and even if there is only one goal of practical reasoning, what an agent has overall reason to do may be difficult to determine.

Philosophers get to play god when they create thought experiments to tap people's intuitions. In such circumstances, it is easy to describe situations where an agent has reason to do only one thing. However, life is complicated, and real agents often have more than one reason for acting. Sometimes the various reasons an agent has for acting will be complementary, reinforcing the degree to which the agent has reason to act in one particular way. If the empty shed is not only burning but also on the point of collapse, the agent may have an even stronger reason to leave. However, reasons may also pull an agent in different directions. If the burning shed contains horses, the agent may have reason to both leave immediately and try to get the animals out.

The complexity of the question of what makes an action one that an agent has overall reason to perform, or one that an agent ought to perform, and the minimal relevance of this to my arguments about justifiable reasons, means that overall reasons are not usually the topic of this thesis. Useful conclusions can be drawn about practical reasons even if it is not overall reasons that are under consideration.

Philosophers use a variety of different terms to refer to the degrees of strength of reasons and the way in which reasons influence or should influence agents. Reasons that favour an action to some degree, but may be outweighed by other considerations and so fail to indicate what a person has overall reason to do, are called 'pro tanto' normative reasons. 'Pro tanto reasons'

---

[5] Patricia Greenspan also argues that an agent can have an overall normative reason to act in some way without it being the case that she ought to act in that way (2007). Greenspan's position is briefly described in the discussion on pro tanto reasons below.

[6] I have been asked whether I take a satisficing or optimising approach to reasons. I find this difficult to answer for two reasons. First, I think that satisficing reasons are usually considered 'good enough' because spending more time to get a better outcome has other costs. Second, although it is easy to think about whether a satisficing or optimising approach is best for overall reasons, I find it difficult to think through the implications, if any, for pro tanto reasons.

is Shelly Kagan's term, with 'pro tanto', meaning 'only to a certain extent' (1989, p. 17). Pro tanto reasons differ from prima facie reasons.[7] He distinguishes between them as follows:

> … a prima facie reason *appears* to be a reason, but may actually not be a reason at all, or may not have weight in all cases it appears to. In contrast, a pro tanto reason is a genuine reason—with actual weight—but may not be a *decisive* one in various cases. (Kagan, 1989, p. 17, italics in original)[8]

Return to the example of someone in a burning shed that contains horses. The agent might be said to have a 'prima facie' reason to save the animals if she is unable to save them, but she would only have reason to save them if she were able to save them. She would have a 'pro tanto' reason to save the horses if she has some reason to save them, but more reason to concentrate on just getting herself out. Reference to prima facie reasons is usually focussed on the situations in which such so-called 'reasons' merely appear to be reasons, or where they would be reasons in other circumstances, but have been overturned by the circumstances. When reasons are referred to as 'pro tanto reasons', only situations in which these reasons have some force as reasons are relevant.

Agents ought to be disposed to take pro tanto normative reasons to favour them acting; that is, if an agent with a pro tanto reason to act in some way were put in a position where she understood her circumstances and reasoned about them appropriately, she ought to be inclined to agree that in her circumstances, the reason favours her acting in a certain way. This is not the same as saying that agents ought to act on pro tanto reasons. Consider the case where pro tanto normative reasons are not outweighed by other reasons, that is, the pro tanto reason is also an overall reason for acting. If you are alone in an empty shed and it catches fire, you ought to take the fire to be a reason to leave. However, even when an agent has overall reason to act in some way, I only assume that this means that the agent ought to be disposed to act in that way. As above, I don't make the stronger claim that an agent with an overall reason to act in some way ought to act in that way. However, agents ought to be disposed to take pro tanto normative reasons to favour them acting even when those reasons are outweighed by other reasons. If the burning shed contains horses and it is a bad thing for animals to die painfully, then the agent ought to be at least somewhat disposed to accept that she has some reason to save the animals, even if she has overall reason to leave immediately.

Patricia Greenspan's so-called 'optional reasons' might be thought a counterexample to my claim that agents ought to be disposed to take pro tanto reasons to favour them acting in some way. 'Optional reasons' are reasons that favour actions, but they are a form of normative reason that agents can ignore without incurring criticism (Greenspan, 2007). Greenspan is correct to claim that some reasons seem to be 'optional', in the sense that it can be fine for an agent to ignore them even though the agent might be thought rational for spending time thinking about them or even acting on them. However, Greenspan says nothing that conflicts

---

[7] Ross uses the term 'prima facie' in a way that is at odds with the use of prima facie given here; using the term as Ross does would entail that a prima facie reason always has force as a reason (Ross, 2003, pp. 19-20).

[8] Torbjörn Tännsjö provides a similar analysis (2010, pp. 29-32).

with my claim that agents ought to be disposed to take pro tanto reasons to favour them acting. Greenspan suggests two ways in which reasons can be optional.

Sometimes a reason favours acting in a certain way, but it favours acting in that way without that entailing that the agent does something wrong if she ignores the reason.[9] Greenspan gives two different examples of this kind of case. In the first example, an agent who looks best in her green blazer decides to wear her blue one instead, because she is in circumstances where looking good is of no concern (Greenspan, 2007, p. 176). In this case, it seems odd to say that the agent ought to be disposed to take herself to have reason to wear the green blazer, but this is because the agent has no pro tanto reason to wear the green blazer. Greenspan describes a situation where the fact that the agent looks better when wearing the green blazer is a prima facie reason, and, moreover, a prima facie reason that doesn't apply in the circumstances described. In Greenspan's second example, an agent holidaying on the Riviera considers visiting Rome (2007, p. 177). The trip to Rome will increase the agent's happiness, but the agent is content to stay where she is. Greenspan argues that the agent does nothing wrong if she chooses to forgo the increased happiness she would get if she travelled to Rome. The Riviera versus Rome example raises different issues. Rather than a case of prima facie reasons, this seems an example of a situation where an agent has an overall reason to act in some way, in this case, to visit Rome, but having an overall reason to act in this way does not mean that the agent ought to act in this way. Even if Greenspan is right, it remains the case that the agent ought to be disposed to acknowledge that she has reason to travel to Rome.

The second way in which Greenspan suggests reasons can be optional occurs when the payoff from practical reasoning falls below a certain threshold (2007, pp. 180-181). Greenspan's examples of this are cases where avoiding certain foods may make the agent's life a tiny bit better or worse. She suggests that in such cases, any benefits may be 'too minor to have to bother with in deliberation, though not so trivial that paying some attention to them would be irrational' (Greenspan, 2007, p. 180). In the case of food additives, the reason isn't considered weighty enough to entail that the agent needs to spend time thinking about it. Greenspan is correct that some reasons matter more than others. There are some reasons that matter so little that saying that an agent ought to be disposed to take them to favour acting in some way will seem excessively demanding. If spending weeks of free time researching food additives and yet more time sorting out which foods to avoid means that you may live for an extra couple of months, should you be disposed to put the required time into the project? The problem with this example is that Greenspan's argument appeals to the costs of practical reasoning, but these costs entail either that the agent has no pro tanto reason to act because the costs of investigating acting are too high, or that the agent has a pro tanto reason to act. The reason that it sounds odd to say that you ought to be disposed to put the time into researching food additives is that for many people the benefits of putting that same time into something else would be significantly greater. If the benefits of putting the time into something else are sufficiently higher, it may be inappropriate for the agent to research food

---

[9] Greenspan calls these 'purely positive reasons', but her analysis of purely positive reasons is not relevant here (2007, pp. 175, 177).

additives and so the agent may have a prima facie rather than pro tanto reason to research food additives.[10] Any agents who do have a pro tanto reason to investigate food additives ought to be disposed to take pro tanto reasons to favour them investigating food additives.

Some reasons matter more than others. Sometimes a reason will matter so much that an agent ought not just be disposed to take the reason to favour acting in a certain way, the agent ought to recognise that the reason favours her acting in that way. Arguably, an agent ought always to take the collapse of an ill stranger next to her to favour her helping him, even if she takes herself to have more reason not to help him. I am not going to consider where the threshold lies between the two kinds of cases.

Jonathan Dancy claims that he uses the term 'contributory reasons' to refer to the same concept of reasons that Kagan refers to as 'pro tanto reasons' (Dancy, 2004, p. 17, fn 13). Dancy calls these 'contributory reasons' because they contribute to what the person has reason to do. The parallels between contributory reasons and pro tanto reasons seem clear. According to Dancy:

> A contributory reason for action is a feature whose presence makes something of a case for acting, but in such a way that the overall case for doing that action can be improved or strengthened by the addition of a second feature playing a similar role. (Dancy, 2004, p. 15)

Contributory reasons are clearly not prima facie reasons, because, in the given context, they really do favour actions to some degree; that is, agents have some reason to act on them even when they have overall reason to act in some incompatible way (Dancy, 2004, pp. 15, 17). In this way, contributory reasons resemble pro tanto reasons. In spite of Dancy's claims and the apparent resemblance between Dancy's contributory reasons and pro tanto reasons, contributory reasons and pro tanto reasons also appear to differ. However, I argue that this difference is only apparent.

The term 'contributory reason' sounds like a reason that contributes to what an agent has reason to do, but is not an overall reason for acting. In an apparent contrast, a pro tanto reason for acting some way may serve as an overall reason for acting in that way. For example, a building's being on fire can give an agent a pro tanto reason to run from the building, and that pro tanto reason can be an overall reason to run. The term 'contributory' sounds as though more than one reason would be needed to contribute to the agent having overall reason to act in that way. If, for example, the building's being on fire is a contributory reason for leaving, it sounds as though something else must also have to be the case for the agent to have an overall reason to leave, such as the building being about to collapse. This connotation of the term 'contributory' is reinforced by Dancy's introduction of what he calls 'decisive reasons'. Dancy's 'decisive reasons' are reasons that operate on their own and mean that the agent ought to take herself to have reason to act in one particular way and no other (2004, pp. 16, 26, 43). Dancy states that decisive reasons are not contributory reasons, because it would be logically impossible for an agent to have a decisive reason to act in one way, but some reason

---

[10] I elaborate on claims such as this in §3.2.

to act in alternative ways (2004, p. 16). This opposition between contributory reasons and decisive reasons might be thought to match a presumed distinction between contributory reasons and overall reasons. However, in spite of the surface appearances, contributory reasons do not combine to make overall reasons, nor does Dancy's distinction between contributory and decisive reasons equate to a distinction between contributory and overall reasons.

In spite of the implications of the name 'contributory reasons', Dancy uses this term to refer to reasons that can also be overall reasons. This claim might seem controversial given that Dancy claims that there are no overall reasons:

> To talk of what there is overall reason to do (and note that 'reason' in this phrase is not a count noun) is to talk about where the contributory reasons come down— on this side or that. We can say that we have more reason to do this than to do that, but most reason to do some third thing. These verdicts do not themselves specify further reasons (of an overall sort), on pain of changing the very situation on which they pass verdict. So there are no overall reasons. All reasons, then are contributory… (Dancy, 2004, p. 16)

Dancy's claims here are less strange than they may seem. When Dancy uses the term 'contributory reasons', the word 'reasons' is a count noun that refers to states of affairs that favour agents acting in certain ways. A burning building is a contributory reason because it is a state of affairs that favours certain actions. Thus, Dancy's contributory reasons are a form of the reasons that favour, or are taken to favour, an action.[11] In contrast, when Dancy refers to 'overall reasons' in the first sentence of the quote, he is referring to what an agent 'has overall reason to do', not to the individual reason or reasons that give the agent reason to act in that way. An agent alone in a burning building has overall reason to run.[12] In spite of Dancy's claim that 'there are no overall reasons', I see no objection to using the term 'overall reason' as a count noun to refer to what Dancy might describe as a contributory reason that overwhelmingly favours an agent acting in a certain way. This is what I did in my discussion of overall reasons at the start of §2.2, and doing so would not conflict with Dancy's position.

In spite of surface appearances, there are no grounds for disbelieving Dancy's claim that his concept of contributory reasons corresponds to Kagan's concept of pro tanto reasons. On Dancy's account, what agents have overall reason to do is determined by the contributory reasons that favour them acting in certain ways, and this is the case even if there is only one so-called 'contributory reason'. Like pro tanto reasons, contributory reasons are reasons that to some degree count in favour of an agent's acting in a certain way. For the sake of simplicity, from here on I use the term 'pro tanto reasons' rather than 'pro tanto or contributory reasons'.

Once Dancy's meaning is clear, his argument that there is no form of reason called an 'overall reason' is convincing. It would be peculiar if there was a distinct category of reasons

---

[11] This claim is explained in more detail in the discussion of what I call 'reason-explanans' in §3.3.

[12] I use the term 'reason relation' to refer to an agent having reason to act in some way. See §3.2.

'overall reasons' that gave agents overall reason to act in some way, especially if one accepts Dancy's arguments that reasons are states of affairs. Pro tanto, prima facie and overall reasons are not different kinds of thing, in the sense that riverbanks, piggy banks and commercial banks are different kinds of things. They differ only in the roles that they play in relation to agents' actions. When we speak about an agent having overall reason to act in some way, we just mean that when the effects of all the pro tanto reasons are considered, that particular action is the one that is most favoured.

In §2.1, I wrote that there can be motivating reasons that never result in an action, that is, pro tanto motivating reasons. It should now be clear what I meant by this claim. Pro tanto motivating reasons are reasons that agents would be disposed to engage with motivationally if they were not outweighed by other considerations. Pro tanto motivating reasons play a role in the mental processing agents engage in when they are deciding what to do, but they need not be the reason on which the agent acted. Not every factor considered when an agent thinks about what she has reason to do will be a pro tanto motivating reason. Pro tanto motivating reasons are factors that the agent took to be reasons for acting, where this is judged independently of whether the agent acted for that reason.

I have two reasons for mentioning pro tanto motivating reasons. First, pointing out another commonality between motivating reasons and normative reasons, namely that both kinds of reasons can take pro tanto form, helps make the parallels between motivating and normative reasons explicit, and this matters for my later arguments. Second, understanding an agent's pro tanto motivating reasons is necessary for understanding and evaluating an agent's actions. It can be useful to understand what things mattered to a person when she was deciding how to act, that is, what aspects of the world were taken to be reasons for acting, or what aspects of the world the agent would have been disposed to take as reasons for acting, not just the reason that the agent assumed she had overall reason to act on. Consider an agent who suffers from weakness of will. Weak-willed Wendy decides that she has most reason to help the person collapsed next to her, but instead pulls out a tissue and blows her nose. Wendy had a pro tanto motivating reason for helping the collapsed man. The way we discuss and evaluate her action would differ from the way in which we would discuss or evaluate the action of a psychopathic agent who had no motivating reason to help the collapsed man. Pro tanto normative reasons are the features of situations that ought to be treated as reasons. Pro tanto motivating reasons are the features of situations that agents took to be reasons.

2.3 Conclusion

The reasons that are primarily of interest in this work are the so-called 'justifiable reasons' that were outlined in the first chapter, but arguing in support of this concept requires situating these reasons within existing philosophical concepts and debates.

My general understanding of the distinction between motivating and normative reasons is more or less standard. I take it that motivating reasons are usually the same kind of thing as

normative reasons, but the category of motivating reasons includes reasons that motivate but are not normatively good reasons. This leads me to interpret motivating reasons a little more broadly than some: I include in this category reasons that have played a part in practical reasoning, but that might not be cited as the cause of any action, perhaps because the agent was prevented from acting as she chose, or because some other reason dominated.

Agents can have some reason to act in a certain way without it being clear that they have most reason to act in that way. The issue of what gives an agent most reason to act is not relevant to this thesis. When I speak of reasons, my intention is usually to refer to what an agent has some reason to do, not what the agent has overall reason to do. Thus, unless I specify otherwise, assume that all talk of reasons here is of pro tanto reasons for action.

Throughout this discussion of motivating, normative, overall, pro tanto and prima facie reasons, I have assumed that the reader knows what is meant by the word 'reason'. However, this word can be far more confusing than the other terms introduced so far. The word 'reason' is ambiguous, and it is this ambiguity that lies behind Dancy's distinction between count noun and non-count noun senses of the word 'reason'. The next chapter distinguishes between these different ways in which the word 'reason' is used.

3. Reasoning, Reasons, and Reason Relations

3.0 Introduction

'Reason' is the most important word used in this thesis, but the word is ambiguous. The word 'reason' can be used to refer to reasoning, reasons, or reason relations. All three of these play vital roles within philosophy of action. Consider the sentence,

> 'I reason to the best of my ability, consider various reasons for acting, and determine what I have reason to do.'

In this sentence, 'I reason' refers to the thought process I go through to sort out what I have reason to believe or do. When 'I consider various reasons for acting', the things that I consider are the reasons that help explain why acting in some way had something to recommend it. Finally, the phrase, 'I have reason to do' refers to a relationship between me as an agent, the world, the action, the intended outcome of that action and, arguably, the reasoning that ties these things together.

This ambiguity of the word 'reason' results in philosophical confusions. Recall Jonathan Dancy's distinction between 'contributory reasons' and 'overall reasons' described in the last chapter (2004, pp. 15-17). This is not a simple distinction between pro tanto and overall reasons. Dancy is also distinguishing between the use of the word 'reason' to refer to reasons that help explain why acting in some way had something to recommend it and the use of the word 'reason' to refer to what agents have reason to do. Both these senses of 'reason' play key roles in the remainder of this thesis, and when the same term is used to refer to both of them it can be difficult to keep clear which 'reason' is at issue. The terms 'pro tanto reasons' and 'overall reasons' cannot be used to resolve this ambiguity, as there can be pro tanto reason relations as well as pro tanto reasons, and overall reasons as well as overall reason relations. To avoid such confusions in the use of the word 'reason', I use the term 'reason-explanans' to refer to the reasons that indicate that acting in some way might be worthwhile, and I use the term 'reason relation' to refer to the 'have reason to' sense of the word 'reason'. I chose the term 'reason-explanans' because I take it that these reasons serve as explanans both in explanations of why someone acted as she did and explanations of what was to be said in favour of someone acting in some way. I use the term 'reason relation' because agents have reason to act in some way only when they are part of a reason relation. When it is unimportant whether I am taken to refer to reason-explanans or reason relations, I simply refer to 'reasons'.

Matjaž Potrč and Vojko Strahovnik claim that 'reason' refers to agents' contexts and it is not appropriate to use 'reason' to refer to facts such as 'he collapsed', which I am calling 'reason-explanans' (2004, pp. 29-34). The fact that someone collapsed is, Potrč and Strahovnik point out, a necessary but not sufficient condition for having reason to help him. It is the agent's context, which includes the fact that someone collapsed, which provides a sufficient

condition for the agent to have a reason to help him. However, I argue below that the word 'reason' can be used to refer both to facts and to the context that exists when an agent has reason to act in some way. This does not belittle the importance of context. As I argue, it is the context – the reason relation – which makes some state of affairs a reason.

This chapter begins, in §3.1, with a brief explanation of what I take reasoning to involve and the importance of reasoning within philosophy of action. Next, I distinguish between reasons and reason relations, and explain what I take such reasons and reason relations to be. In §3.2 I discuss reason relations. I have a relational theory of reasons, that is, I hold that for something to be a reason, it needs to be part of a reason relation. This emphasis on reason relations differs from the work of philosophers such as Derek Parfit and R. Jay Wallace (Parfit, 1997, 2001; Wallace, 2003a). Even many philosophers who treat reasons as relational fail to explicitly acknowledge that they are doing so or spend little time discussing the nature of the relation.[13] In §3.3 I discuss reasons themselves, that is, 'reason-explanans', and I discuss an objection to the position on motivating reasons I described in §2.1.

3.1 Reasoning

Practical reasoning is necessarily tied to action. I take the word 'action' to refer to the intentional behaviour of an agent, that is, I assume that actions as such cannot be accidental. Assume that someone collapses next to you and you move towards him, intending to offer help. If you trip, fall, and break his arm, you broke his arm unintentionally. In such a situation, your breaking of his arm is not an action because it is not an intentional response to reasons.

Intentional actions may result from conscious deliberation or from unconscious mental processing that it would be strange to call 'deliberation'. When an agent researches a situation, reflects on the advantages and disadvantages of acting in various ways, and consciously decides that she has most reason to act in one particular way she engages in conscious deliberation. Alternatively, conscious deliberation might involve briefly weighing up the situation in a quick and less sophisticated way. Agents can, however, act intentionally without engaging in such intellectualised reasoning.[14] An agent may decide to act in a certain way without ever consciously thinking about how she should act. An experienced nurse, for example, may automatically help someone who appears to be ailing without reasoning about how it would be best for her to act or even forming a conscious belief about the person's ailment (Benner, 2001). In such situations, the lack of conscious deliberation may be a sign of the agent's skilfulness in responding to reasons rather than an indication that the agent was not responding to reasons. It may seem peculiar to call such unconscious mental processing 'practical reasoning', but it is directed at action, it does involve reasons, and the reasoning required can be reconstructed and evaluated. My aim here is to consider reasons in a way that does not presume that all intentional actions follow from conscious deliberation.

---

[13] See, for example, (Adams, 1993; Dancy, 2000, 2004, 2009; Joyce, 2001; Scanlon, 1998; Setiya, 2004).

[14] Karen Jones and Peter Railton argue for this position, and Ralph Wedgwood provides a sketch of an argument against Kieran Setiya's rejection of this view (Jones, 2003; Railton, 2009; Setiya, 2007; Wedgwood, 2008).

The elements and processes involved in practical reasoning are of more interest in this thesis than the performance of the action, even though the performance of an action is the goal of practical reasoning.[15] As in the case of the person who trips and falls, an agent may engage in practical reasoning and form an intention to act, yet be unable to carry out that action due to factors that are outside of her control. Practical reasoning can also take place a long time before any action occurs, or after an action has been performed. When an agent reflects on an action that she has already performed, or continues to deliberate about an action she has already decided to perform, she is still engaging in practical reasoning.

For some philosophers, particularly those with Kantian sympathies, appropriate reasoning is an important part of what determines whether an agent has a normative reason to act in some way (Korsgaard, 1986; Smith, 1994, 1995a, 1995b, 1996). Philosophers with more consequentialist intuitions may hold that reasoning has little to do with what people have reason to do, because what people have reason to do is either determined by facts about surrounding circumstances and the outcome of actions, or determined by the expected utility of an action (Hawthorne & Stanley, 2008, p. 571; Tännsjö, 2010). These distinctions become clearer in Chapter 8.

3.2 Reason relations

The reason relation that exists when an agent has reason to act in some way differs from the reasons that give an agent reason to act in some way. However, non-philosophers and philosophers use the same word 'reason' to refer to both reason relations and reasons, and this obscures the distinction. The difference between Dancy's concepts of 'contributory reasons' and 'overall reasons' described in §2.2 is just one instance of this confusing ambiguity. The reasons that give an agent reason to act in some way explain why the agent has, or thinks she has, reason to act in that way. If the building that I am in is on fire, the fact that the building is on fire is part of the explanation of why I have reason to leave the building. So when it is not clear whether I am using 'reason' to refer to reasons or reason relations, I call them 'reason-explanans'. Something is a reason-explanans only when it is part of a reason relation. A building can be on fire without that giving me or anyone else reason to do anything. What makes the building's being on fire a reason for me to act is that I am in it, and that my dying a painful death is (I hope) undesirable on any account of desirability. These things together mean that I have reason to leave the burning building. This dependence of reason-explanans on reason relations means that reason relations are the key to understanding reason-explanans and practical reasoning. I am primarily interested in the nature of reason relations in philosophy of action.

---

[15] Thomas Scanlon discusses the output of practical reason in terms of intentions rather than actions (1998, pp. 20-22). This has the advantage of emphasising that the formation of an appropriate attitude is more important than the performance of the action. However, it is simpler to talk of agents having reasons to act, rather than agents having reasons to form intentions to act.

I use 'reason relation' to refer to the conditions under which an agent has or had reason to act in some way. There are two senses in which an agent can have reason to act in some way.[16] On the first reading of the phrase, the agent has reason for acting in that way because she knows that certain facts favour her acting in that way. For example, if you want to reduce your alcohol intake, and you know that cider and gin are available and that gin has more alcohol than cider, you have reason to drink cider rather than gin. There is a reason for you to act in a certain way and you 'have it' because you know that reason-relation applies to someone in your situation. This is not the sense of 'having reason to act' that is of interest here. When I write that an agent 'has reason to act in some way', then, unless I specify that I am referring to motivating reason relations, I mean only that the agent's situation is such that some state of affairs favours her acting in some way. The agent may never 'have' the reason in the sense that she knows that some reason relation applies to her, and the agent may have reason to act in a way that differs from the way that she takes herself to have reason to act.

A normative reason relation statement is true when it is at least part of an explanation of why an agent ought to be at least somewhat inclined to act in some way. For any agent A, action φ, set of circumstances C, and reason(s) R:

> Positive normative reason relation: A has reason to φ if and only if, in C, R favours A's φ-ing.[17]

In the reason relation, 'A has reason to φ', the word 'reason' is, in Dancy's words, not a count noun (2004, p. 16). It refers to the reason relation of which A is a part. When 'A has reason to φ' is true, it may be true that there is a reason, in the count noun sense, for A to φ, but it may be a complex of things that give the agent reason to act rather than a single thing. To say that there is a reason for A to φ, is to say that if someone asks, 'Why does A have reason to φ?' it makes sense to pick out one aspect of A's circumstances and point to that as an explanation of why A has reason to φ. Which aspect of A's circumstances is picked out as the reason that explains why A has reason to φ will depend on the circumstances and intent of the person who asks the question.

To see the difference between reason relations and reason-explanans, consider the following example. I have reason to fly to Wellington. The set of things that lie behind my having reason to fly to Wellington include things to do with me, things to do with the set of circumstances that surround me, values (I will explain what I take this to mean later), and

---

[16] Mark Schroeder makes this distinction clear, but when he does so, he doesn't distinguish between an agent having a reason to act and an agent having reason to act (2008b). I discuss this distinction below.

[17] A similar formulation can be used to express a situation where an agent has no reason to act in a particular way:

> Negative normative reason relation: No reason favours A's φ-ing if and only if, in C, there is no reason that favours A's φ-ing.

This expression of a negative normative reason relation can be adapted in ways that correspond to the various formulations of positive reason relations throughout this thesis. However, negative reason relations play no part in arguments about reasons given here, so negative reason relation formulations are not given.

many rules of inference. They include such things as my appointment to a new position, my eagerness to meet the people I will be working with, the importance of attending work when required to do so, the first meeting being in Wellington, the administrator's buying me a ticket to fly there, and so on. If someone asks me what reason I have to fly to Wellington, which one of the many possible explanations I present as 'the reason' that favours me acting in that way will depend on who asks the question. If someone I work with in Christchurch asks why I have reason to fly to Wellington, the reason that best explains what favours me acting in that way may be that I have recently been appointed to a new position. If a family member who knows about my new job asks what reason I have to fly to Wellington, the correct reason to mention may be that my first meeting will be held there. Agents have reason to act in certain ways because of a complex of their characteristics and circumstances. To say that an agent is part of a reason relation, that is, to say that the agent 'has reason to act' does not point to any aspect of the agent's circumstances that serves as a reason.

As I argued in Chapter 2, motivating and normative reason relations differ slightly, but have the same general form. A motivating reason relation statement is true when it is at least part of an explanation of why an agent is, or was, somewhat motivated to act in a certain way.

> Positive motivating reason relation: A has reason to φ if and only if, in C, A took R to favour φ-ing.

The circumstances in which A took herself to have reason to act do not play the same role that they do in normative reasons, but 'C' is left in for consistency's sake, and because the agent's circumstances will affect any judgement made about the normative appropriateness of her action.

The term 'favour' is a somewhat awkward description of the link between reason-explanans and actions, but it seems to be the word that is most often used in philosophical discussion of reasons (Crisp, 2006, pp. 19, 37-39; Cullity & Gaut, 1997, p. 1; Dancy, 2000, pp. 1-2; Scanlon, 1998, pp. 17-19; Schroeder, 2008c, p. 11). The term 'favour' might be taken to imply that there is some person who takes one thing to provide evidence in support of something else. However, where reasons are concerned this is not necessarily the case.

Saying that something 'favours' an action may mean that someone takes it to favour an action, or it may mean that it is the kind of thing that would be taken to favour the action if certain conditions were met. The first of these situations fits motivating reasons well. If an agent takes something to be a reason for acting, and responds by acting, then that presumed reason served as a motivating reason for the agent's action, and the agent took it to favour her acting as she did. But, there are two ways in which this use of 'favour' does not fit well with normative reasons. First, recall that normative reasons are good reasons for agents to act, but agents do not always identify what there is good reason for them to do. An agent can have a normative reason for acting some way without accepting that some state of affairs favours her acting in that way. However, if an agent has a normative reason to act some way, then the reason favours her acting in that way in the second sense of the word, that is, the reason is the kind of thing that would be taken to favour the action if certain conditions were met. The

exact nature of these conditions is arguable; those familiar with arguments about normative reasons may immediately think of debates between internalists and externalists about the conditions that must obtain for an agent to have reason to act, but the nature of the conditions that must obtain for an agent to have reason to act are also important to my arguments about justifiable reasons and to other arguments about reasons discussed in this thesis.[18] The relevant point here is just that a reason can favour an agent's acting some way in a counterfactual sense even when no one has identified it as a reason. Second, a situation might be taken to favour an action being carried out in the abstract, that is, independently of there being an agent to carry out the action. On this way of understanding the 'favouring' involved in reason relations, if someone collapses and there is nobody there to help, it might still be said that his collapse is a reason for him to be helped. This way of understanding reasons for action as stemming from circumstances independently of the existence of agents that can act on those reasons, or would be willing to act on those reasons, is relevant to arguments about the nature of objective normative reasons, and it is discussed in Chapter 8.

John Skorupski and Mark Schroeder also distinguish between reasons and reason relations, although their accounts differ slightly from mine. According to Skorupski, reason relations are five-place relations that specify the degree to which facts give agents reason to act in a particular way at a particular time. Skorupski writes:

> [Propositions about reasons] are about *reason* relations…. *Practical* normative propositions are about reasons to act: *the fact that* p *gives* x *reason of degree* d *at time* t *to* Ψ, where Ψ ranges over action types…. For example, the fact that the building is about to explode gives you very good reason to get out right now. (Skorupski, 1997, pp. 345-346, italics in original)

Similarly, almost ten years later, he writes:

> Reasons are facts, or as we can represent it, sets of facts, which stand in a certain relation to an agent and an act-type—also a degree of strength of the reason and a time. Thus we have:
> $R(p_i, x, c, d, t)$
> The set of facts p_i gives x reason of degree d at time t to c. (Skorupski, 2006, p. 28)

The facts or sets of facts that Skorupski refers to as 'p' and 'p_i' and that he describes as 'giving' agents reason to act are what Dancy and many others refer to as 'reasons' and what I refer to as 'reason-explanans'. My focus is on what it means for agents to have reason to act rather than on what 'gives' agents reason to act, and this is one way in which Skorupski's account differs from mine. An agent has reason to act in some way when her name can be substituted for the variable 'x' in Skorupski's analysis of practical normative propositions.

Skorupski is right that reason relations have differing degrees of strength and that they apply at particular times, but I have not included these variables in my account of reasons. The strengths of reasons are not relevant to my thesis about justifiable reasons, so I have not added degrees of strength into my reason relation. As Schroeder points out, calculating the

---

[18] Debates about internalism and externalism about reasons are discussed in Chapter 7.

degrees of strength of reasons is complicated by the possible incommensurability of reasons and the ways in which reasons interact with each other (2008c, pp. 124-126). Rather than mention degrees of strength, I specify that I am not intending to argue about overall reasons. Rather than make specific times relevant to whether agents have reasons to act in particular ways, I include agents' circumstances in the reason relation, as my arguments in support of justifiable reasons rely on the fact that agents' reasons for acting are affected by their circumstances. Anything that makes a particular time relevant to what the agent has reason to do will be captured by the agent's set of circumstances.

Mark Schroeder also discusses reasons for action in terms of reason relations (2008b, 2008c).[19] Schroeder dislikes what he describes as Skorupski's 'four-place reason relation', because he dislikes Skorupski's inclusion of the strength or weight of reasons (Schroeder, 2008c, p. 101, fn 27). Instead, Schroeder introduces a three-place reason relation:

> *R* is a reason for *X* to do *A*. (See, for example, Schroeder, 2008c, pp. 15-17, 207)

After arguing for his Humean theory of reasons, he explains this reason relation as follows:

> For *R* to be a reason for *X* to do *A* is for there to be some *p* such that *X* has a desire whose object is *p*, and the truth of *R* is part of what explains why *X*'s doing *A* promotes *p*. (Schroeder, 2008c, pp. 59, 193)

Consider Schroeder's account of reason relations using the collapsed person example. For the collapse of the person next to you to be a reason for you to go to his aid is for there to be a character trait of beneficence such that you have a desire to be beneficent, and the fact that the man collapsed is part of what explains why if you were to go to his aid, your action would exhibit beneficence. This is a complicated way of saying that his collapse gave you reason to help him because you want to be a kind person. I agree with Schroeder that what it means to say that a reason favours an agent's acting in some way can be explained in more detail, but my way of understanding reason relations differs from Schroeder's.

To decide what an agent has reason to do we seem to need to sort out four things: who the agent is; the state of the world; what has value; and ways of getting what has value. Thus, reason relations are relationships between agents, actions, and the world, including those things that serve as reason-explanans. My use of the word 'value' here is about as deflationary as it can get. It carries no commitment to any particular theory of value, nor does it imply that the agent has a strong commitment to bringing about an outcome. In brief, an agent has a motivating reason to act when she believes that if she carries out some action, then, given her situation and the general state of the world, her action will bring about an outcome that she values. An agent has a normative reason to act when some state of the world is a sign that acting in a certain way will produce a certain result, the result has some value, and the value of achieving that result is not significantly outweighed by conflicting values.

---

[19] John Broome states that normative requirements are relations, and also refers to the 'reasons relation' (Broome, 1999, p. 400).

This way of understanding reason relations may seem to imply that practical reasoning only involves determining appropriate means for achieving our goals, but this is an overly simple interpretation of my account of reason relations. My reference to 'goals' refers to more than just the states of affairs that may result from actions. The goal of an action, in this broader sense of the word, may be that it exhibits a certain kind of character, rather than that it has a certain kind of result. Second, practical reasoning is also involved in working out what goals are appropriate. People may need to consider interactions between many of the things that matter to them to determine their goals, and considering the relationships between the things that currently matter may lead people to set new goals or realise that something that they cared deeply about is not really important.

An agent only has reason to act when the agent is in circumstances such that carrying out some action will produce a certain end, the end has some value, and that value is not significantly outweighed by the negative value produced by that same action. Return to the example where you are travelling on a bus and the person next to you collapses. For you to have normative reason to help the collapsed person, it needs to be the case that:

(1)	The person collapses.	A state of affairs.	*STATE*	S
(2)	The collapse is a sign that trying to help the person might benefit him.	If a particular state of affairs exists, then if a particular action is carried out, a particular effect will follow.	*END*	$S \rightarrow (\varphi \rightarrow E)$ 'φ' = action 'E' = effect
(3)	Benefitting the collapsed person has some positive value.	If a particular effect occurs, positive value will be realised.	*VAL*	$E \rightarrow V$ 'V' = positive value
(4)	Trying to help the collapsed person would realise that value without conflicting with other values to such a degree that there is no reason to try to help him.	If it is the case that if a certain action occurs, it will realise certain negative values, then it is not the case that those negative values significantly outweigh the positive values of acting in that way.	*SIG*	$(\varphi \rightarrow N) \rightarrow \neg(N > sigV)$ 'N' = negative value '$>sig$' = significantly greater

(1) is usually taken to be a state of affairs, proposition, or belief about the world. I refer to it as '*STATE*', for reasons explained later.

(2) states, in effect, that given a state of affairs, an action will yield an effect; it gives the relationship between the state of affairs, an action, and the

effect of the action. I call it '*END*' because, broadly speaking, it expresses a means-end relationship.

(3) states that the effect has positive value; it gives a prima facie reason for someone to be at least somewhat motivated to carry out the action. I label it '*VAL*' for value.

(4) states that the value of the effect of the action is not significantly outweighed by possible negative values that may result from the same action; that is, the prima facie reason for carrying out that action is not significantly outweighed by other reasons. I label it '*SIG*' because it says that the value of carrying out the action remains significant.

The reason relation that this yields remains a pro tanto, rather than overall, reason relation. If the man next to you on the bus has collapsed, then, if conditions *END, VAL*, and *SIG* obtain, his collapse favours your helping him. This does not mean that you have overall reason to help him, just that you have a pro tanto reason to do so. The mere collapse of a person is neither a motivating nor a normative reason for you to act, nor would a mere belief that he collapsed or that he was ill serve as a motivating or normative reason.[20]

My description of (1) to (4) sounds as though they occur sequentially: first, the person collapses; second, the agent reasons from the collapse to the possibility that a particular action will have a particular result; third and fourth, the result of the action is evaluated. However, *STATE, END, VAL*, and *SIG* may occur in any order. Someone may start out the day contemplating the importance of carrying out random acts of kindness and imagining situations where random acts of kindness are well- or ill-advised, that is, she might start out the day thinking abstractly about situations that resemble *SIG*. When she gets on the bus she is already searching for ways to improve someone's functioning, and looking for someone who she might have reason to help. In this sequence of events, *STATE*, the state of affairs that I am calling the agent's reason, may be the last event in the process of practical reasoning.

In this thesis, I usually take *STATE* to be a state of affairs that serves as the reason-explanans for both motivating and normative reasons. Sometimes when *STATE* is treated as the agent's reason, it is taken to express a proposition or a psychological state, usually a belief, rather than a state of affairs.[21] For example, an agent's motivating reason for helping someone might be taken to be that she believed that he had collapsed. Similarly, the agent's normative reason for helping might be taken to be the truth of the proposition that he collapsed. For the sake of simplicity, I always treat *STATE* as though it is a single, relatively uncomplicated state of affairs, but sometimes it is a combination of interacting states of affairs. In these circumstances '*STATE*' will need to be treated as referring to the complex rather than to one component of that complex. I explain why I take *STATE* to be the reason-explanans and a state of affairs rather than a proposition or a psychological state in §3.3.

---

[20] In other words, I agree with Potrč and Strahovnik's claim that reason-explanans are necessary, but not sufficient, for an agent to have reason to act (2004).

[21] Something like this distinction in ways of thinking about reasons is also noted by Alfred Mele (2007, p. 100).

STATE is just one part of the explanation of why an agent has, or had, reason to act in a particular way. Because it is just one part of the explanation, other components of the reason relation, *END, VAL, SIG, E, V*, and *¬(N>sigV)* are also sometimes treated as the agent's reason for acting. I see no grounds for claiming that it is inappropriate to call any of these different elements of the reason relation 'the reason' in the reason-explanans sense of the term; that is, any of them could be taken to be the '*R*' in the positive normative reason relation '*A* has reason to *φ* if and only if *A* is in *C*, and *R* favours *A*'s *φ*-ing'. Whichever of *STATE, END, VAL, SIG, E, V*, and *¬(N>sigV)* are not treated as *R* get subsumed within the set of circumstances, *C*.

END involves reasoning to a conclusion about the implications of acting in a certain way given the existence of the state of affairs expressed by *STATE*. This conclusion is about the outcome of an action in the broadest sense of the word 'outcome'. As I wrote above, on this sense of the word, the relevant 'outcome' may be that the action is an act of a certain kind rather than that the action has certain consequences. If, for example, you attend a party when you feel like staying home, the relevant outcome may be that you showed loyalty to a friend, and this may be the relevant outcome even if your friend was unaware that your attendance was an act of loyalty, or even unaware that you attended the party. As with *STATE*, I describe the effects of actions in simple terms, but *END* can stand in for a complex of effects that only give an agent reason to act because they exist within that complex.

The description of *END* may sound as though I am assuming that relatively sophisticated conscious reasoning is required for someone to determine that she has a reason to act. As I argued in §3.1, this is not the case. In practice, the state of affairs, *STATE*, and the implications of acting in certain ways given that state of affairs, could be so tightly connected that the reasoning in *END* is automatic rather than the result of conscious deliberation. This sometimes leads to the formation of a belief, in this case, a belief that helping the person might improve his functioning. However, the reasoning in *END* may be the basis for the agent's action even if it does not lead to the agent forming a conscious belief. Thus, an agent may move to help without ever consciously thinking, 'Helping the person may improve his functioning'.[22]

Like *STATE*, *END* can be treated as a psychological state or a proposition that favoured the agent's acting in a certain way. The belief formed through the processing described in *END* may be treated as the motivating reason for an agent's action. One plausible answer to the question, 'What reason did she have for helping him?' is that she believed that her helping him would improve his functioning. This can be treated as the agent's motivating reason even if she had no conscious, or even dispositional, belief that helping him would benefit him.

---

[22] My use of *END* in the reason relation is not the same as Stephen Finlay's claim that reasons are 'end-relational' because 'a fact is a reason for *φ*-ing, relative to a system of ends E, iff it explains why *φ*-ing is conducive to E…. [Where by an 'end' he means] a possible aim for action or object of desire' (Finlay, 2006, p. 8). Finlay's 'E' seems to be composed of attributes from all of *END*, *VAL*, and *SIG*. I do, however, like Finlay's way of picking out why a state of affairs would count as a reason-explanans in the '*STATE*' sense of the word.

Perhaps she acted automatically, but if she were questioned afterwards she would have said that her reason for helping him was that she believed it would improve his functioning. Alternatively, the reasoning in *END* might be treated as a proposition that serves as an agent's normative reason for acting:

> 'Why does she have reason to help him?'

> 'If she were to help him, it would improve his functioning.'

As stated above, although the beliefs or propositions that result from the processing in *END* can be described as agents' reasons, they are not what I take to be the reason-explanans in this thesis.

VAL connects the likely outcome of the proposed action and the value of that outcome. If the potential outcome of the action didn't matter at all, the action would not be worth doing. As with *STATE* and *END*, *VAL* is sometimes treated as a reason in the sense of being the factor that explains why the agent has reason for acting. And, as with *STATE* and *END*, *VAL* may be treated as a statement about an agent's psychological states or a statement about an objective state of affairs. For instance, *VAL* may be treated as a desire. It might be claimed that the agent's motivating reason for helping the collapsed man is that she cared about him. Similarly, the agent might be said to have a normative reason to help him because she cares about him. If *VAL* is treated as a statement about the objective value of an action, the agent might be said to have helped him or to have reason to help him because helping people is a good thing to do.

The introduction of *VAL* may suggest a conflict between my account of reasons and various accounts of the relationship between desires, values, and reasons. My talk of values and goals may, for example, seem to conflict with Thomas Scanlon's rejection of what he calls 'teleological' approaches to reasons, or with his so-called 'buck-passing' account of value (1998, pp. 79-103, 108). Use of the word 'value' may also be taken to conflict with Humean theories of reasons. However, my aim is to give an account of reason relations that avoids taking any particular position on what it means to say that something has value or on the relationship between reasons, values and desires.

On Scanlon's buck-passing account of values, to say that something has value just is to say that it has some property that gives agents reasons to act in certain ways (1998, p. 97). This need not conflict with my use of the word 'value', as my use of 'value' can be read as a claim that something has a property that makes it worth doing or bringing about. So, those who accept the buck-passing account of value can substitute a buck-passing friendly phrase for my phrase 'either has, or is taken to have, some value.' For instance, someone who holds a buck-passing view can substitute for *VAL* the statement:

> *VAL* [buck-passing]. Improving the collapsed person's functioning either has, or is taken to have, some property that makes it worth doing.

If desired, Scanlon's alternative interpretations of what it means to say something has value, such as that it is appropriate to do it, worthy of respect, or fitting to do it, could also be

substituted for my use of the word 'value' (1998, pp. 95, 98). Similarly, '*V*' could be taken to refer to a property that makes something worth doing and '*N*' to a property that makes it worth not doing. Thus my use of the word 'value' can be interpreted in such a way that what it means for some state to be valuable is consistent with Scanlon's buck-passing account.

This buck-passing interpretation of *VAL* might still be taken to conflict with Scanlon's account of reasons, as it could be read, in a teleological way, as a claim that there is reason to bring about a certain state of affairs, in this example, the improvement of the collapsed person's functioning (1998, p. 98). However, my basic account of reason relations is compatible with Scanlon's rejection of teleological interpretations of reasons (1998, pp. 79-94, 108). As above, the so-called 'outcome' of an action may be that an act of a particular kind was carried out rather than that an act had a certain consequence. Nor do I claim that whenever something is taken to be valuable, that particular thing is to be promoted. *SIG* is introduced because I do not assume that this is the case. My basic account is also consistent with Scanlon's pluralism about values (1998, p. 108). As I wrote in §2.2, I intend to leave open the possibility that practical reasoning serves many functions, and different functions could be associated with different values.

I also aim to avoid making claims that commit me to, or conflict with, so-called 'Humean theories of reasons'. Broadly speaking, a Humean theory of reasons assumes that what an agent has reason to do is explained by that agent's psychological states, usually the agent's desires, but possibly by what the agent finds pleasurable or thinks valuable.[23] A Humean about reasons might claim that *VAL* only makes sense if 'the outcome of the action either has, or is taken to have, some value' is read as a claim that the agent desires, or would desire, the outcome of the action. In contrast, someone who rejects Humeanism about reasons might claim that *VAL* only makes sense if 'value' is taken to refer to agent-independent objective values. Because the way in which the word 'value' is understood has no implications for my overall thesis about reasons, I am happy for 'value' to be read as referring to either desires or to objective values independent of any agent's psychological states.

The final component of reason relations, *SIG*, is important because, as mentioned in §2.2, I want to discuss pro tanto or overall reasons rather than prima facie reasons. Sometimes carrying out an action would have a valuable outcome in some circumstances, but no value in other circumstances. Assume, for example, that the bus is on fire, the collapsed person is highly likely to die whatever the agent does, and the agent is likely to die unless she gets out of the bus fast. In this case, any value associated with helping the collapsed man conflicts with the value of getting out of the bus fast, and the weight of the values is such that the agent has no reason to help the collapsed man. The agent would, therefore, have a prima facie, rather than pro tanto, reason to help the collapsed man.

As with the other components of reason relations, *SIG* can be treated as the agent's reason for acting. The agent might be said to have had reason to help the collapsed man because she

---

[23] Schroeder discusses several possible ways of developing a Humean theory of reasons (2007).

wanted to help him more than she wanted to do anything else, or to have had reason to help him because helping him would not have negative consequences that overwhelmed the value of helping him. And, as these examples show, *SIG*, like *VAL*, can be taken to depend on an agent's experiencing a particular psychological state, such as desiring an outcome, or *SIG* can be taken to be a function of the psychological state that it is most appropriate for the agent to have. *SIG* can also be interpreted as a statement about values. As above, the particular interpretation of *SIG* is not relevant to the principle arguments in this thesis.

Hallvard Lillehammer's contrast between accounts of reasons that take reasons to be 'given by facts' and those which take reasons to be 'given by ends' offers unintentional support to my analysis of reason relations (2010, p. 17). In his article, Lillehammer considers the possibility that there is a conflict between so-called 'fact-based' and 'end-based' accounts of normative reasons. He gives many examples of philosophers who he suggests take one or the other approach to reasons. His examples of facts-based reasons include works by Derek Parfit and Jonathan Dancy (Dancy, 2000, pp. 69-70; Parfit, 1997, p. 130; cited in Lillehammer, 2010, pp. 18-19) His ends-based reasons are examples of reasons that are based on values of some kind; they include works by Parfit, Michael Smith, and Christine Korsgaard (Korsgaard, 1996, pp. 115-116; Parfit, 1997, p. 130; Smith, 2004, pp. 60-61: cited in Lillehammer, 2010, pp. 18-19). After considering the roles that facts and ends play in accounts of reasons, Lillehammer concludes that there is little to be said in favour of taking a fact-based approach to understanding reasons rather than an end-based approach, or vice versa (2010, p. 17). If my account of reasons is correct, then Lillehammer's conclusion is to be expected. If I am right, any correct account of normative reasons will have components that are both fact- and end-based; there is no conflict here, just a difference in the researchers' focus.

The distinction just drawn between *STATE, END, VAL*, and *SIG* enables the assumptions in my initial reason relation formulation to be teased apart in a way that helps illustrate the nature of, and distinction between, reason relations and reason-explanans. According to my positive normative reason relation:

> *A* has reason to φ if and only if, in *C*, *R* (reason-explanans) favours *A*'s φ-ing.

However, as explained, what *C* and *R* are taken to be will depend on the circumstances, interests and intent of the person asking for an explanation of what the agent has reason to do. The elements of the reason relation that it would not be appropriate to treat as reason-explanans, given the focus of the inquiry, get bundled into '*C*'. However, the reason relation can be expanded to make its components more explicit:

> *A* has reason to φ iff for *A* in *C*, *S*, $S \rightarrow (\varphi \rightarrow E)$, $E \rightarrow V$, and
> $(\varphi \rightarrow N) \rightarrow \neg (N > sigV)$.

What it means for something to be a reason-explanans is expressed by the reason relation as a whole; that is, the reason-explanans is a reason because of its relationship to the other components of the reason relation. If *S* is a reason-explanans, this will be because of its relation to $S \rightarrow (\varphi \rightarrow E)$, $E \rightarrow V$, and $(\varphi \rightarrow N) \rightarrow \neg (N > sigV)$.

When people write and speak of 'reasons for action', they are often referring to the existence of the reason relation itself rather than identifying or making claims about an agent or any other components of the reason relation. In such circumstances, they might say that 'there is a reason to act' rather than '*A* has a reason to act'. When people use 'has a reason' in this way, they are not referring to a reason that some agent somehow 'has', whether the 'has' is taken to mean that the agent knows about the reason or whether the 'has' is taken to mean that there is a reason that gives a particular agent reason to act.[24] Instead they are claiming that there is a state of the world such that some action would realise a certain significant value.

There is reason to φ iff S, $S{\rightarrow}(\varphi{\rightarrow}E)$, $E{\rightarrow}V$, and $(\varphi{\rightarrow}N){\rightarrow}\neg(N{>}sigV)$.

This could also be represented as a counterfactual claim that if there were an agent in a position to carry out some action, the agent would have reason to act in that way. This kind of claim can be made for many reasons. For example, it may be a first step in an argument that everyone has reason to act in a certain way, or in an argument that claims that people have reason to put themselves in a position where they can act in a certain way in the future.

The versatility of this way of describing reason relations can be seen by the way in which this account can be used to understand Dancy's distinction between contributory, decisive and overall reasons, and to understand arguments about whether desires are reasons.[25] The reason-explanans that is used to explain why the agent has reason to act in a certain way can be picked out in several ways. When philosophers argue about whether desires are reasons they often treat the potential value of an action as the explanation of why the agent has reason to act. In such situations, the reason, *R*, is substituted for the expression of positive value, *V*:

R is a reason for *A* in *C* to φ iff for *A* in *C*, *S*, $S{\rightarrow}(\varphi{\rightarrow}E)$, $E{\rightarrow}R$, and $(\varphi{\rightarrow}N){\rightarrow}\neg(N{>}sigR)$.

Dancy's contributory reasons correspond to pro tanto reasons, but Dancy's specific focus is on reasons of the kind described by *STATE*. If the situation calls for an explanation of what the agent has reason to do in terms of *STATE*, the reason, *R*, is substituted for the state of affairs, *S*:

R is a reason for *A* in *C* to φ iff for *A* in *C*, *R*, $R{\rightarrow}(\varphi{\rightarrow}E)$, $E{\rightarrow}V$, and $(\varphi{\rightarrow}N){\rightarrow}\neg(N{>}sigV)$.

Dancy uses the term 'decisive reason' to refer to reason-explanans that operate on their own and mean that the agent ought to take herself to have reason to act in one particular way and no other (2004, pp. 16, 26, 43). In other words, decisive reasons are reasons that cannot be outweighed by any possible consequences of the agent acting in that way.

R is a reason for *A* in *C* to φ iff for *A* in *C*, *R*, $R{\rightarrow}(\varphi{\rightarrow}E)$, $E{\rightarrow}V$, and $(\varphi{\rightarrow}N){\rightarrow}\neg(N{>}sigV)$.

When Dancy writes that 'reason' in the phrase 'talk of what there is overall reason to do' is 'not a count noun', what he means is that to talk of what there is overall reason to do is to

---

[24] This contrasts with the position Schroeder argues for in 'Having Reasons' (2008b).

[25] Dancy's distinction between these concepts of reasons was introduced in §2.2.

refer to the reason relation as a whole, not to pick out some element of the reason relation that is the overall reason-explanans (2004, p. 16).

Debates about reasons tend to be split into arguments about two different issues, each of which involves arguments about the nature of what are taken to be reason-explanans. The first set of debates is over the use of the word 'reason' to refer to *STATE* or *END*. This set of debates tends to focus on how the way the world is, or the way the world is taken to be, affects the actions that agents have reason to carry out. Arguments related to *STATE* and *END* usually focus on whether reason-explanans are beliefs, propositions, or states of affairs. The second set of debates about reasons includes arguments about the issues raised by *VAL* and *SIG*, including arguments over whether desires are ever reasons, whether values can be reasons, or whether desires are a necessary background condition for having reason to act. The focus here is on debates of the first kind.

3.3 Reasons

When I refer to 'reason-explanans' in the remainder of this thesis, I use the term to refer to *STATE*, and I assume that reason-explanans are states of affairs and not, in the normal case, propositions or beliefs. (This does not conflict with my claim that *END, VAL*, and *SIG* can all be treated as reason-explanans, it is just that this is not my focus here.) I write 'in the normal case' because there are occasions when propositions and beliefs serve as de facto states of affairs. This is explained in more detail later. In treating reason-explanans as states of affairs, I follow philosophers whose general approach to philosophy of action differs from mine, such as Robert Audi, Jay Wallace, and Jonathan Dancy (Audi, 2001; Dancy, 2000, 2009; Wallace, 2003a). In what follows, I first explain why I assume that reason-explanans are states of affairs rather than propositions. Afterwards, I use one of Dancy's arguments to show that there are grounds for holding that reason-explanans are states of affairs rather than psychological states. Like Dancy, I take it that motivating and normative reason-explanans are the same kind of thing; that is, both normative and motivating reason-explanans are states of affairs (Dancy, 2000). However, I argue that Dancy draws mistaken conclusions about the implications of his argument, and explain the way in which my conclusions differ from those of Dancy.

Philosophers who believe that there are good arguments against reason-explanans normally being psychological states, tend to claim that reasons must be either facts, propositions, or states of affairs. I do not discuss the possibility that reasons are facts rather than states of affairs, because the meaning of the word 'fact' is unclear, usually involves reference to states of affairs, propositions, or both, and I want to avoid this debate (Lemos, 1995, pp. 15-19; Loux, 2002).

When writing about reasons, it is easier to treat them as true propositions than it is to treat them as states of affairs. Schroeder decides to assume that reasons are propositions because 'things go more smoothly if we say that reasons are true propositions than if we say that they

are facts' (2008c, p. 21). I assume that Schroeder uses the word 'fact' to refer to states of affairs, as he distinguishes between facts and true propositions, and his examples of facts serving as reasons include such things as, 'Obesity is a reason to eat well' (Schroeder, 2008c, p. 20). Propositions lend themselves nicely to inclusion in arguments that use standard philosophical systems of logic, and this is probably what Schroeder means by 'things go more smoothly'. However, states of affairs are what people, including philosophers, usually take to be reasons.[26] It isn't the proposition 'that the shed is on fire' that is a reason to evacuate the shed. The reason to leave is the shed's being on fire. The world provides us with reasons, not our statements about the world.

It is unlikely that anything of great importance for my account of justifiable reasons follows from whether reasons are propositions or states of affairs. Although I claim that reasons are states of affairs, my theory of reasons could be tweaked to refer to propositions rather than states of affairs. If reasons are states of affairs, then the shed's being on fire is a reason for you to leave. If reasons are propositions, your reason for leaving would be that the shed was on fire. Schroeder seems to agree that nothing of importance follows from this, as he follows his statement that he will take reasons to be propositions with the claim that it doesn't matter much whether reasons are taken to be true propositions or facts (2008c, p. 21).

Philosophers who argue against accounts of reasons that resemble my account of justifiable reasons tend to assume that accounts like mine are based on an assumption that reason-explanans are psychological states. When accounts like mine are taken to be based on psychological states, they are described as 'subjective' and often dismissed for this reason.[27] My opponents' assumption is false; it is easy to argue for my overall position while taking reasons to be states of affairs or propositions rather than beliefs or other psychological states, and I aim to show that this is the case. I also want to ensure that my arguments about the nature of reason relations deal with reason-explanans in a way that is as similar as possible to the approach of the people with whom I disagree. Jonathan Dancy, Jay Wallace and Mark Schroeder, for example, argue that neither motivating nor normative reasons are psychological states (Dancy, 2000; Schroeder, 2008c; Wallace, 2003b; 2006). However, if motivating and normative reasons were shown to be psychological states, this would not undermine my thesis that justifiable reasons are a viable and important form of normative reasons for action. If reason-explanans were shown to be psychological states, then the arguments given by Dancy, Wallace and Schroeder that make my position on justifiable reasons seem initially implausible, would have been overturned. Arguing in support of justifiable reasons in this situation would be significantly easier.

The most obvious objection to my claim that motivating and normative reasons are usually states of affairs rather than psychological states is that while normative reasons may not be psychological states, motivating reasons must be psychological states. Michael Smith holds

---

[26] This is one part of Jonathan Dancy's argument that reasons are not psychological states (2000). The argument is explained in more detail below.

[27] This issue is considered in more detail in Chapter 6.

this view (1994, p. 96). Assume that a stranger fakes a collapse next to you and you decide to help him. I claim that your motivating reason was a state of affairs rather than a psychological state. This may seem implausible. Motivating reasons explain actions. You might think that if you had not believed that he was ill, you would not have acted as you did. If this is so, your false belief that he was ill seems to have been the motivating reason that explains your action. The alternative, saying that your reason for acting was a state of affairs, may seem crazy. What illness? He was faking!

Jonathan Dancy argues against this claim that motivating reason-explanans are psychological states, and his argument seems correct (2000, pp. 121-137). Dancy points out that when people are asked why they acted in some way, they usually mention a state of affairs, not their psychological states. If, as you move to help the collapsed stranger, I ask you what reason you have for doing it, you are likely to say, 'He collapsed'. If I ask you the same question later, when you discover that the stranger was faking, you may mention your psychological states, and say instead, 'I believed he was ill'. So, when people accept that the reason they had for acting was a good reason for acting, they usually describe that reason as a state of affairs, but when people discover that they were somehow mistaken in their understanding of what there was reason to do, they alter their response, saying instead that they only *believed* that state of affairs to be the case. In other words, people usually cite states of affairs as reasons, apart from when they learn they were mistaken. Dancy argues that one account of reasons is better than two; that is, it is preferable to have an account of motivating reasons that functions both in situations where people's beliefs are correct and in those where their beliefs are false. It also seems desirable to be able to say that the same kind of reason-explanans can be both motivating and normative. Yet, the prevailing opinion is that normative reasons are not psychological states. So, it would be convenient if an account of motivating reasons could be given on which they were also not psychological states. According to Dancy, such an account is possible. I agree with him that it is possible, but give a different account from Dancy of why it is possible.

Dancy argues that explanations of agents' reasons for action can be non-factive, that is, the state of affairs given as an explanation of why the agent had reason to act need not exist. Dancy draws an analogy with false convictions. Convicts serve prison terms for carrying out crimes. Someone falsely convicted of murder is still imprisoned for murder. If we know that she didn't commit murder, then we might explain the reason for her imprisonment by saying that she allegedly committed murder. In a similar fashion, if you moved to help the person who collapsed next to you, but he was faking, we might say that you moved to help him because he was ill, even though he wasn't. If it sounds odd to say that you moved to help him because he was ill when he wasn't ill, then Dancy suggests that some equivalent of 'allegedly' is added to the statement that he was ill. To quote Dancy:

> One may feel some unease about saying 'His reason was that *p*' when we don't ourselves believe that *p*. But this … should not drive us to saying 'His reason was that he believed that *p*', supposing that by this device we are respecting the factive nature of explanation; we can avoid any apparent commitment on our part

to things being as the agent supposed by use of one of a number of special constructions such as 'as he supposed'. (Dancy, 2000: 136)

So, on Dancy's account, motivating reasons are the states of affairs that agents took to favour them acting as they did. This will lead to clumsy statements about non-factive states of affairs when someone knows that things were not as the agent supposed. However, Dancy believes that this accords with our standard practices. The alternative, as Dancy sees it, is to claim that all motivating reasons are psychological states, and this, he argues, leads to more serious difficulties for theories of reasons.

Although Dancy uses words that are not terms for psychological states in his non-factive explanations, this change in terminology seems nothing more than an elliptical way of referring to psychological states. The claim that what explains the convict's imprisonment is that she 'allegedly committed murder', seems equivalent to saying that a group of people in the justice system believed that she committed murder rather than equivalent to saying that the explanation for her imprisonment was the state of affairs 'she committed murder'. Similarly, the claim that your reason for helping the person who faked an illness by collapsing was that you supposed that he was ill, seems another way of saying that you helped him because you believed that he was ill. Thus, even though Dancy is correct that people say such things, he is not correct to claim that this avoids introducing beliefs into explanations of agents' reasons for acting. I agree with Dancy's claim that it would be best to have an understanding of reason-explanans on which they are a single kind of thing whether they are motivating or normative. I also agree with him that reasons are most naturally thought of as states of affairs. However, I don't believe that it is necessary to introduce the notion of non-factive explanations to achieve this goal. Dancy's approach, like that of many others, picks out the wrong states of affairs as reasons.

If the faker's illness is identified as the agent's reason for helping him, it is difficult to see how the agent's reason for acting could be a state of affairs. The error is in taking the agent's reason to be the collapsed person's illness. If you were to help the person collapsed next to you, your reason for helping him would be his collapse. The collapse is the event in the world that spurred your reaction. His collapse may have led you to reason to the conclusion that he was ill, and you may have formed a belief that he was ill, but it is still his collapse, which is a state of affairs, that spurred your reasoning and your action.[28] My approach achieves Dancy's aims; it fits with agents' tendencies to mention states of affairs, rather than beliefs, when asked why they acted. It also enables reason-explanans to be, metaphysically speaking, the same thing whether they are motivating or normative. Nevertheless, my account does not require non-factive explanations of reasons.

---

[28] As explained in §3.1, I do not mean to suggest that people sit down and engage in reasoning about the contents of the world whenever they are motivated to act. You may have acted before you had a chance to form a belief that he was ill. If an action results from a disposition to aid those who appear to be in trouble, there seems no reason to doubt that it is an intentional action.

Here is another example of the difference between what I treat as reasons, in the precondition reason-explanans sense, and what some other philosophers treat as reasons. Bernard Williams refers to reasons in the way in which I wish to refer to them. When Williams presents the case of a man who wants a gin and tonic, but has at hand a bottle containing petrol rather than gin, he asks whether the man has reason to pour the 'contents of the bottle' into the glass, mix it with tonic and drink it (1981a). Williams does not ask whether the man has reason to pour the alleged gin into a glass, mix it with tonic, and drink it. The gin drinker may believe that the bottle contains gin. He would be disposed to say that the bottle contains gin if he were asked about its contents. However, if he believes that it contains gin, this is the result of reasoning based on the state of the world: in this case the bottle containing the liquid and his general situation. Perhaps he is handed a bottle labelled gin while at a party. Jonathan Dancy, in contrast, treats reason explanations as non-factive (2000, pp. 131-135, 146-147). So, according to Dancy, even when there is petrol in the gin bottle it makes sense to say: 'The man poured the contents of the bottle into a glass, mixed it with tonic and drank it because, allegedly, by doing so he would make a gin and tonic.' Similarly, according to Schroeder, it makes sense to say the reason that the man has to take a sip from his glass is that 'his glass contains gin and tonic', even though it contains petrol and tonic (2008c, p. 15 fn 23:2). I favour Williams' approach over that of Dancy and Schroeder. The man's reason for pouring the stuff into a glass and drinking it is not that the bottle allegedly contains gin, or that he believes it contains gin; it is that the stuff is in a bottle labelled 'gin' sitting on a table at a party.

Although I have argued that reasons are not psychological states, it is important to note that psychological states can be reasons. The fact that we believe, desire, hope, fear, or feel anything is not sufficient for having a motivating or normative reason to act in accordance with that belief, desire, and so on. However, psychological states can be treated as states of affairs that serve as reasons within reason relations.

To see the distinction between the contents of psychological states and treating psychological states and capacities as states of affairs, consider the situation of someone who is furious. If someone is furious, her psychological state can be treated as a state of affairs that affects what she has reason to do. For example, if she knows that when she is furious she is likely to say things that she will regret later, her fury can be a state of affairs that justifies her leaving the room. The agent's reason in such a case would be the state of 'being furious'. This differs from someone storming out of the room in a raging fury, where it is questionable whether the person is acting as an agent at all. It also differs from the situation of someone yelling furiously, where it may be appropriate to take the agent's reason for yelling to be an action carried out by someone else rather than her fury.

Dancy and Schroeder also distinguish between treating psychological capacities and psychological states as states of affairs and treating them as sets of subjective experiences. Dancy agrees that psychological states can be reason-explanans, but argues that reason-explanans are states of affairs. (See, for example, 2000, p. 157). Depression, for example, can be both a psychological state and a state of affairs. Depressed people find it difficult to

motivate themselves to do anything. Although the lack of motivation due to their depression may not give them reason to do nothing, it does give them reason to seek help from a doctor or psychologist. Mark Schroeder makes a similar point. He writes that 'facts about what an agent believes may themselves be objective normative reasons for her to act. For example, the fact that Jon believes that he is Napoleon may be an objective normative reason for him to seek psychological counselling' (2008c, p. 14 , fn 19, spelling altered). This suggests that psychological capacities and states may be treated as states of affairs, even though they are mind-dependent, and also treated as affecting the reasoning and reason-explanans on which reason relations are based without introducing worrying subjectivity.

3.4 Conclusion

The ambiguity of the term 'reason' means that the terms 'normative reasons' and 'motivating reasons' are used ambiguously. Sometimes 'reason' is used to refer to the things that indicate or explain why agents have, or take themselves to have, reason to act in some way. However, 'reason' can also be used to refer to reason relations, that is, to agents having normative or motivating reasons for acting. In the rest of the thesis, when I describe states of affairs as reasons, I refer to reason-explanans. When I speak of an agent having reason to act, I refer to a reason relation. When it matters whether I am referring to a reason that explains an action or what an agent has reason to do, and it might be unclear which I am referring to, I use the terms 'reason-explanans' and 'reason relation'. I take it that whether reason-explanans are motivating or normative, they are usually states of affairs, but what makes those states of affairs reasons is that they have a particular kind relationship with agents or other circumstances. Different conceptions of normative reasons follow from different assumptions about the nature of such reason relations.

The claim that reasons are psychological states or propositions conflicts with my claim that reasons are states of affairs. However, if I were wrong, this would not undermine my arguments, given in the next two chapters, that justifiable reasons are a viable and important form of normative reason. If reasons were conclusively shown to be psychological states, it would be easier to argue in support of justifiable reasons. If reasons were shown to be propositions rather than states of affairs, only a few minor changes would need to be made to my arguments in support of justifiable reasons.

4. Justifiable Reasons

4.0 Introduction

Justifiable reasons are reasons that ordinary people ought to take to be pro tanto reasons for acting. In this way, they contrast strongly with motivating reasons and objective normative reasons. An agent can have a motivating reason for acting in some way even though she ought not to take herself to have reason to act in that way; perhaps she negligently failed to consider her options. Similarly, an agent can have an objective normative reason for acting even though she ought not to take herself to have reason to act in that way; perhaps she would need to act in bizarre and inappropriate ways to learn she had reason to act in that way. I argue that an agent has a justifiable reason for acting some way if and only if, were the agent to consider the circumstances in a way that is practically and epistemically appropriate, she would hold that some state of affairs favours her acting in that way. The requirement that agents consider their circumstances in a way that is appropriate for agents in their circumstances is the source of normativity for justifiable reasons. Because my account of justifiable reasons only requires agents to reason in a way that is practically and epistemically possible and appropriate for them, rather than possible and appropriate for an ideal observer, we can speak of what agents with a wide range of intellectual and physical abilities have justifiable reason to do. Remarkably little has been written about reasons that resemble my account of justifiable reasons. Philosophers developing accounts of practical reason tend to focus on accounts of motivating reasons and objective normative reasons; that is, they tend to focus on what people took themselves to have reason to do or on what people in highly idealised situations ought to take themselves to have reason to do. Motivating and objective normative reasons matter, but are not a useful guide to real people's normative reasons for acting.

Motivating reasons are not a guide to what people ought to take themselves or others to have reason to do; they provide a particular type of description of what an agent took to be a reason for acting. People recognise that what they take themselves to have reason to do is not always the same as what they ought to have taken themselves to have reason to do. Sometimes people's reasonings and actions seem to be exactly as they ought to be. Sometimes people reason and act in ways that they are proud of, no matter the confusions in their reasoning or outcomes of their actions. Sometimes people take themselves to have reason to act in ways that they later realise were foolish. Talk of motivating reasons is talk of what happened, or what contributed to what happened, not an evaluation of an agent's reasons for acting.[29]

Objective normative reasons are intended to be a guide to what agents have objective reason to do; they are directed at correct outcomes or values, where 'value' is interpreted broadly. Accounts of such reasons play an important role in theorising about ethics, values, and practical

---

[29] Motivating reasons are discussed more fully in Chapters 2 and 3.

reason, and I discuss their importance in Chapter 8. However, according to accounts of objective normative reasons, what agents have reason to do will often be something that they cannot learn they have reason to do Accounts of objective normative reasons require people to do such things as ensure that they have a coherent set of beliefs and desires, ensure that they know all the facts relevant to their situation, or ensure that they know the consequences of possible actions and the value of those consequences.[30] Such accounts of reasons describe practical reasons that real human agents cannot always use as a basis for action.

The possibility of developing an account of normative reasons that gain their normativity from ideals of justification set by human capacities has generally been overlooked or rejected by those who research practical reason. A few people have sketched out accounts of subjective normative reasons, but the accounts given vary significantly, and I argue in Chapter 6 that they do not provide an account of the reasons that ordinary people ought to take to be reasons for acting.[31] I want to provide an account of practical reasons that it is possible and appropriate for real human agents to act on, rather than a description of the reasons that motivated people's actions or an account of practical reason for ideal agents.

Practical reason at the level of ordinary human agents matters. I argue that justifiable reasons are the only normative reasons on which human agents can base their actions; any real human agent who acts for a normative reason, whether that be a moral or prudential reason, acts for a justifiable reason. In everyday life, evaluation of the appropriateness of someone acting in a certain way for a certain reason is tied to issues of self-regulation, rationality, praise and blame. I argue that whether a reason is a good reason for a normal human agent to act depends on whether someone would be justified in taking it to be a good reason for acting. The 'someone' who I describe as 'a normal human agent' may have very superior intelligence or borderline intelligence, although the justifiable reason relations that apply to those with superior intelligence and borderline intelligence may sometimes differ because of their differing reasoning abilities.[32] These someones may also have varying physical skills and know-how, and, again, this will affect what they have justifiable reason to do. The conceptions of the normative reasons that people take account of in their day-to-day lives serve a valuable function, and if these concepts are coherent, it is appropriate for philosophers to treat them as functional and valuable concepts. I argue for the consistency and usefulness of a form of reasons, 'justifiable reasons', which corresponds to real human reasons for acting.

---

[30] There are many accounts of reasons that take this approach. Some of these are discussed later in Chapter 8. But, see, for example, the accounts of reasons given by Jonathan Dancy, Derek Parfit, Thomas Scanlon, Kieran Setiya, Michael Smith, Jay Wallace, and Bernard Williams (Dancy, 2000, 2004; Parfit, 2001; Scanlon, 1998; Setiya, 2004; Smith, 1994; Wallace, 2003a; Williams, 1981a).

[31] Some epistemologists have developed accounts of reasons that come closer to mine by arguing that the correct account of theoretical reasons will also apply to practical reasons (Fantl & McGrath, 2009). This approach to understanding practical reasons is not discussed here. Errol Lord develops an account of 'factoring reasons' which makes some claims that are very similar to mine (2010). His account is described in more detail in §4.4.

[32] Someone with 'Very Superior Intelligence' will have an IQ score of over 130. Someone with 'Borderline Intelligence' will have an IQ of between 70 and 79 (Wechsler, 2002).

Return to the example from the previous chapters: You are on a bus, worrying that your stuffy nose means that you have a head cold. Just before your stop, the person next to you collapses. Although you cannot know this, he is faking. If you help him, he will continue to act in this way, which will cause many more upsets than benefits. What do you have reason to do? You have no objective reason to help the collapsed person. If you knew all the relevant information and reasoned appropriately, you would realise that he did not need help. If consequences determine what you have objective reason to do, better consequences follow from not helping him. However, you are likely to think that if someone collapses next to you, you have reason to help. My account of justifiable reasons claims just that. The concept of justifiable reasons is normative, and it successfully accounts for situations where real agents' reasons for acting fail to correspond to ideal agents' reasons for acting. Thus, the concept of justifiable reasons is well placed to serve as an interface between talk of reasons and appropriate self-regulation, praise, blame and rationality.

I begin, in §4.1, by briefly describing the advantage of justifiable reasons over objective normative reasons for explaining our common sense understandings of self-regulation, praise, blame and rationality. In §4.2 I present my concept of justifiable reasons and reason relations in more detail. A reason for acting is only a justifiable reason if an agent can learn that she has reason to act in that way after engaging in possible and appropriate practical reasoning, so in §4.3, I clarify what I take such possible and appropriate deliberation to be. In §4.4, I describe the relationship between justifiable reasons and other forms of reasons, and explain that although justifiable reasons are not reducible to motivating or objective normative reasons, they are not different in kind from motivating reasons or objective normative reasons.

4.1 Justifiable reasons, self-regulation, rationality, praise and blame

Appropriate self-regulation, rationality, and praiseworthy behaviour are commonly taken to involve responding appropriately to pro tanto and overall normative reasons. So, the link between reasons and self-regulation, rationality, praise and blame is both useful and intuitively plausible, but it is a link that concepts of objective normative reasons cannot explain. In contrast, self-regulation, rationality, praise and blame can all be explained with reference to justifiable reason relations. In this way, justifiable reason relations serve a function that objective normative reason relations cannot serve.

Self-regulation involves agents deciding how they will act, rather than letting the world, their impulses, or other people carry them along.[33] So, self-regulating agents engage in practical reasoning. Self-regulating agents make decisions in response to internal or external factors and, unless something intervenes, act on those decisions. In the absence of good arguments for doing otherwise, it seems appropriate to call the internal and external factors that real people consider or act on when regulating their behaviour the 'reasons' for their actions. Self-regulation is a normative process; agents can regulate their own behaviour in right or wrong, and good or bad, ways. So, unless there are good grounds for thinking otherwise, the internal

---

[33] As explained in §3.1, 'decide' does not mean 'decide after conscious deliberation'.

and external factors that real people ought to consider or act on when they regulate their behaviour are normative reasons for their actions.

Assume that appropriate self-regulation could involve two different sorts of normative reasons, objective normative reasons or justifiable reasons. Self-regulation cannot always involve objective normative reasons, because objective normative reasons are not always appropriately accessible to real agents. Consider a nurse who has been trained, for good reason, to avoid expressing disgust when faced with various injuries and excretions. Nurse Tuckett is caring for a heavy smoker with stinking skin, clothes and breath. If he were to reason to the best of his ability, he would conclude that he has reason to suppress his disgust to avoid causing offence or distress. But, in this particular case, if Nurse Tuckett had the knowledge and reasoning ability of an agent with a gods'-eye perspective, he would learn that expressing disgust would encourage the patient to give up smoking. In this situation, what Nurse Tuckett has objective normative reason to do cannot play a role in his self-regulation. As things stand, he simply cannot access the information necessary to learn what he has objective normative reason to do. Even if we assume that there is some form of investigation Nurse Tuckett could carry out that would reveal that showing his disgust for this patient's state would be a good thing, it would not be appropriate for him to delay caring for the patient to carry out such investigations. But, self-regulation is still called for, because we expect people in non-ideal practical and epistemic situations to engage in appropriate self-regulation. In contrast, the concept of justifiable reason relations that I develop is a form of accessible normative reasons that can always be used for self-regulation.

Judgements about rationality within everyday practical reasoning are determined by considering what an agent would take herself to have reason to do if she were to reason and act to the best of her ability.[34] Real human agents are not considered practically irrational for failing to see that they have reason to act in a way that only agents with a gods'-eye perspective would be able to see there was reason to act. If an internationally renowned cancer specialist tells Nurse Tuckett to give the smoker a drug, and nothing suggests that the specialist has made an error, Nurse Tuckett would act rationally if he gave the patient the drug. He would be acting rationally even if the patient was dangerously allergic to that drug. If he decided not to give the patient the drug, he would be acting irrationally, even if he had no reason to give the patient the drug from a gods'-eye perspective. Given that real agents are often unable to learn what they have objective normative reason to do, judgements about rationality in everyday practical reasoning cannot assume that agents ought to act in response to objective normative reasons. Derek Parfit pointed this out fifteen years ago (Parfit, 1997, 99). In contrast, justifiable reasons dovetail perfectly with judgements of rationality. An agent acts rationally when he acts appropriately in response to states of affairs that it is possible and appropriate for him to take to be reasons. The relationship between reasons and rationality is discussed again briefly in §7.2.2, and it is discussed more fully in Chapter 9, where I argue in support of this claim that an agent acts rationally when he acts appropriately in response to states of affairs that it is possible and appropriate for him to take to be reasons.

---

[34] Errol Lord makes a similar claim (2010). Lord's position is discussed in §4.4.

Appropriate praise and blame are also related to agents' justifiable reasons for acting rather than to objective normative reasons. There is nothing odd or incoherent about blaming agents for failing to see that they have reason to act in some way, or praising them when they see that they have reason to act in some way. However, it would be inappropriate to blame an agent for failing to see that he has reason to act in some way when he could not learn that he has reason to act in that way. If his training and reasoning were justifiable, Nurse Tuckett would be praiseworthy if he suppressed his disgust when faced with a stinking cancer patient. He would not be thought blameworthy for being unable to see that he had reason to express his disgust. What an agent has objective normative reason to do is often inaccessible to real human agents. So when agents are praised or blamed for recognizing what they have reason to do, they are not praised or blamed for recognizing what they have objective normative reason to do.

The link between praiseworthiness and justifiable reasons and the disconnection between praiseworthiness and objective normative reason applies to pro tanto reasons and overall reasons, and moral reasons and prudential reasons. It would be inappropriate to blame an agent for failing to see that he has reason to save a drowning child when he is deaf and cannot hear the child's cries. But, a child drowning is an objective normative reason to save the child, although it might only be a pro tanto reason if the child is Hitler. An agent who wants to get to a shop fast has a prudential objective normative reason to take the quickest route. But, it would be inappropriate to blame him for not taking the quickest route if he had no reason to suspect that the shortest route was blocked by road works. An agent is not morally blameworthy for failing to determine that he has reason to act some way when only an agent with a gods'-eye perspective could determine there was reason to act in that way.

The relationship between justifiable reasons and self-regulation, rationality, and praiseworthiness will seem to some to establish that my position is flawed, especially given that I use this relationship to argue for the importance of justifiable reasons. Those who develop theories of objective normative reasons accept that it is inappropriate to blame people for failing to recognize that they have reason to act in some way when it would be unreasonable to expect them to see that they have reason to act in that way (Smith, 1997b, 2003; Williams, 1995a, 40). Some of those who argue in support of objective normative reasons claim that issues of self-regulation, praise, blame and rationality are distinct from issues relating to reasons (Dancy, 2009, 106-109; Hawthorne & Stanley, 2008, p. 586; Parfit, 1997, 2001; Wallace, 2003a; 2007, pp. 188-189). However, those who respond to the conflict between rationality, praiseworthiness and conceptions of objective normative reasons by suggesting that rationality, praise and blame are not directly related to normative reasons do so because they believe that there cannot be a useful normative concept of what I refer to as 'justifiable reasons'. On the face of it, if such an account could work, it would have the advantage of supporting intuitive links between these concepts. Their rejection of the idea that there can be a concept of normative reasons that is integral to our concepts of self-regulation, rationality and praise and blame is a response to anticipated flaws in approaches

like mine rather than a view for which they argue. The most significant response to these arguments is to show the usefulness and coherence of the concept of justifiable reasons.

4.2 What are justifiable reasons and reason relations?

Justifiable reason-explanans are states of affairs that an agent would take as a sign that there is something to be said in favour of acting in a certain way if her reasoning was appropriate given her practical and epistemic circumstances. They are normative reasons because they are the states of affairs that the agent would be justified in treating as reason-explanans if the agent's reasoning was appropriate given her circumstances. The accessibility of such reasons is vital to their status as reasons; it must be appropriate to expect individual agents in particular circumstances to work out that there is something to be said in favour of them acting in the relevant way. In contrast, accounts of objective normative reasons typically treat the accessibility of such reasons to real human agents as irrelevant to whether the agent has reason to act in that way.[35] They also typically ignore the appropriateness of an agent's taking a certain state of affairs to be a reason for acting when determining whether that state of affairs is a normative reason for acting.

Justifiable reason-explanans are reason-explanans because of their role within justifiable reason relations. This is just to say that the account of normative reasons given in Chapter 3 also applies to justifiable reasons. The justifiable reason relations that apply to agents are determined by considering the reasoning that it is epistemically and practically possible and appropriate to expect from the agents. The initial social, physical, and psychological circumstances of an agent are the starting points and limiting conditions for determining what that agent has reason to do. Other agents will only have justifiable reason to act in the same way if their circumstances are relevantly similar. What an agent with high intelligence has justifiable reason to do may differ from what an agent with low intelligence has justifiable reason to do in what are otherwise identical circumstances. Agents cannot have justifiable reasons to act in ways that they cannot learn they have reason to act. Any agent is required to reason about her situation in a way that would be possible given her circumstances. Agents are also required to reason in a way that would be appropriate given their circumstances. This requirement that agents' reasoning or deliberation be appropriate sets the standard that ensures that justifiable reasons are normative. What this reasoning amounts to is discussed in more detail below. Agents' best possible understandings of the consequences of actions will usually be vital for determining what those agents have justifiable reason to do. However, the actual consequences of any possible actions that agents might perform are less important to what those agents have reason to do than the process of deliberation. If an agent deliberates about the circumstances to the best of her ability and her action results in a bizarrely unpredictable outcome, this is not a sign that she failed to respond appropriately to her justifiable reasons for acting. Similarly, if the agent fails to deliberate appropriately and, as a consequence, acts in a way that brings about the best result, she did not act as she had

---

[35] I argue for this claim in §8.1. Jonathan Dancy is an exception to this rule. His position is discussed in §5.3.2.

justifiable reason to act. These requirements mean that it is both possible and appropriate for agents to use justifiable reason relations to guide their actions.

Justifiable reason relations are a form of normative reason relation. In §3.2, I defined normative reason relations as follows. For any agent, *A*; action, *φ*; set of circumstances, *C*; and reason, *R*:

> Positive normative reason relation: *A* has reason to *φ* in *C* if and only if, in *C*, *R* favours *A*'s *φ*-ing.

As I wrote above, agents have a justifiable reason to act in some way when the information available to them is such that if they were to reason in a way that would be both possible and appropriate for them, they would conclude that they have at least some reason to act in that way. Consequently:

> Positive justifiable reason relation: *A* has justifiable reason to *φ* in *C* if and only if, were *A*, in *C*, to reason in a way that is possible and appropriate, she would hold that *R* favours her *φ*-ing.[36]

The reason relation is what matters most here; my account of justifiable reasons and reason relations is a holistic account. It is the relationship between the agent and the agent's reasoning abilities and circumstances that determines which states of affairs will function as reasons for the agent to act, and determines what the agent will have justifiable normative reason to do.

As with other forms of normative reason relations, what is meant by 'reason' and 'circumstances' can be unpacked further. An agent has reason to act in some way when there is a state of affairs, such that, given that state of affairs, an action will have an effect, the effect has positive value, and the value of that effect remains significant in spite of any negative values produced by the action. Thus, for any state of affairs, *S*; action, *φ*; effect, *E*; positive value, *V*; negative value, *N*; and where '$>sig$' means significantly greater:

> *A* has justifiable reason to *φ* in *C* if and only if, were *A*, in *C*, to reason in a way that is possible and appropriate, she would be aware of *S*, and conclude that $S{\rightarrow}(\varphi{\rightarrow}E)$, $E{\rightarrow}V$, and $(\varphi{\rightarrow}N){\rightarrow}\neg(N{>}sigV)$.

As emphasised in §3.2, my use of the word 'value' is not intended to commit my account of justifiable reasons to any particular theory of values or to any position on debates about Humean theories of reasons. 'Value' may, for instance, be read in a way that is consistent with or conflicts with Scanlon's buck-passing account of reasons without affecting my overall theory (1998).

The first variable in my formulation of reason relations is always *A*, for agent. This is partly because my principal interest is in developing an account of normative reasons that, unlike

---

[36] Negative justifiable reason relation: No justifiable reason favours *A*'s *φ*-ing in *C* if and only if, were *A*, in *C*, to reason in a way that is possible and appropriate, she would hold that no reason favours her *φ*-ing.

objective normative reasons, can always be detected by agents. I hold that what an agent has justifiable reason to do is always constrained by the agent's nature and her circumstances. This means that in a certain sense of the words, justifiable reasons are always agent relative.[37] This does not mean that there are no actions that all normal human agents have justifiable reason to do or not do. Kant may have been right to argue that there are certain actions that agents have reason to do simply in virtue of their agency. Perhaps all agents always have reason not to treat other agents merely as means to their ends (Kant, [1785] 1993, p. 39). Nor does it mean that we cannot speak of what there is justifiable reason to do in situations where there is no agent who can carry out the action. If a child is drowning and nobody is nearby, the reason relation can be expressed counterfactually, for instance, were an agent in circumstances where it would be possible and appropriate for her to learn she had reason to help, the agent would have reason to help. Such counterfactual normative reasons have important roles within philosophy, and justifiable reasons, as much as objective normative reasons, can play such roles.

The normativity of justifiable reasons stems from their justifiability. The term 'normative reason' is used to refer to good reasons for acting. As I argue in §2.2, if a reason is normative, then an agent ought to be disposed to take it to favour her acting in a certain way. Similarly, if a reason relation is normative, an agent ought to be disposed to take it that she has at least pro tanto reason to act in that way. When I claim that justifiable reasons are normative, I am claiming that when it is possible and appropriate for an agent to take her circumstances to favour her acting in a certain way, she ought to be disposed to take the circumstances to favour her acting in that way. Similarly, if an agent's circumstances are such that it would not be possible and appropriate for her to take herself to have reason to act some way, she ought not to be disposed to take the circumstances to favour her acting in that way. What it means to say that it would be possible and appropriate for an agent to take herself to have reason to act in some way is explained in more detail below. The key point, however, is that when it is possible and appropriate for an agent to take herself to have reason to act in some way, the agent would be justified in taking herself to have at least some reason to act in that way, and unjustified were she not to take herself to have reason to act in that way. Although I leave the question of the normativity of justifiable reasons for now, I consider more arguments against the normativity of justifiable reasons in later chapters.

4.3 Possible and appropriate practical reasoning[38]

Justifiable reason relations are the product of possible and appropriate practical reasoning, which means that this concept of reasons relies on intuitions about the relationship between possibility, appropriateness and justification. The reasoning that it is possible and appropriate for agents to carry out when deciding how they will act is affected by agents' situations as human beings living in particular environments at particular times; what an agent has

---

[37] In Chapter 7, I argue that this agent relativity of justifiable reasons is not worrying.

[38] Some of this section is taken, with minor alterations, from (Mason, 2006).

justifiable reason to do is determined by the information and reasoning practically accessible to that agent, rather than all the available information, perfect reasoning, or the actual consequences of any action. Agents' practical and epistemic circumstances affect the information that it is appropriate to expect them to access and the reasoning that it is appropriate to expect from them, and, thus, what they have justifiable reason to do.

My theory of justifiable reasons relies on a vague, intuitive notion of appropriateness. This vagueness arises partly from complexity and from the vagaries of inductive reasoning, but it also arises from uncertainties about what it is possible and appropriate for agents to do.

Used as I use it here, the word 'appropriate' can be taken to require different things by different people in different situations. I accept this, and embrace it as part of my theory. What someone would be justified in taking to be appropriate practical reasoning will be affected by that person's circumstances, which will affect what that person claims there is justifiable reason to do. However, any specific claim that it is appropriate for an agent to take herself to have normative reason to act in some way must itself be supported by good reasons, and those reasons must include an argument that shows that it was in practice possible for the agent to access the information that marked the action out as being one that she had reason to perform.

What it means for it to be possible for agents to correct their beliefs and actions is vague because there are degrees of possibility. At one extreme, an agent can do something because it is possible for him to do it with little mental or physical effort. At the other extreme, an agent cannot do something because doing it is logically impossible. There is a grey area between 'immediately doable' and 'logically impossible' where it is difficult to determine whether someone can or cannot do something. When someone says: 'A can come to know that he has reason to φ', it may mean that the agent can learn that he has reason to act in some way with a little effort, or it may mean that the agent is ignorant of some fact that it would be extremely difficult, or in practice impossible, for him to learn.[39] This vagueness means that it will often not be possible to establish whether an agent has a justifiable reason to act in some way. However, like Williams, I think that this is an advantage of any theory of normative reasons (Williams, 2001). Sometimes what we have reason to do is obscure, and it can be hard to work out whether we really had or have reason to carry out some action. Whether an agent has a justifiable reason to act in some way may even at times be indeterminate.

The vague intuitive notion of appropriateness and possibility on which my account of justifiable reasons relies does not make my concept of justifiable reasons less clear than concepts of objective normative reasons. Analyses of objective normative reasons refer to ideals such as 'all relevant truths', 'sound deliberation' and 'coherent sets of beliefs and desires'. Accounts of objective normative reasons that are based on the notion of relevant

---

[39] This idea that 'could' comes in degrees differs from Smith's black and white treatment of 'could' in 'Rational Capacities' (2003b). But, this less clear-cut notion of 'could' is compatible with Smith's rational capacity approach.

truths rely on an intuitive understanding of what it takes for a truth to be relevant to what an agent has reason to do. Williams' notion of sound deliberation is deliberately left vague to better fit with human projects and capabilities (1981a, p. 104). Similarly, Smith argues that it is not possible to give an explicit analysis of the reasoning required by his account of objective normative reasons (1994, pp. 161-164). So, I take it that the notions of appropriateness and possibility I explain are no less developed than the assumptions used to support objective normative reasons.

Our understanding of what it is possible for agents to do rests on our understanding of when it is appropriate for agents to further investigate what they have reason to do. It is stupid and inappropriate for an agent to learn he has reason to act in some way when it would be wildly inappropriate for him to carry out further investigations into what he has reason to do.

People have an intuitive notion of when it is possible and appropriate for agents to take themselves to have reason to act in some way. Darwin Awards are given to people who 'ensure the long-term survival of our species by removing themselves from the gene pool in a sublimely idiotic fashion' (Northcutt, 2007, p. 2). Reports of Darwin Award winners provide examples of people who failed to learn what they had justifiable reason to do, although they usually blur the distinction between pro tanto and overall reasons. For example, the man who arrived at hospital with a suspected 'sexual proclivity for animals' and 'his entire crotch area…filled with porcupine quills' seems to have failed to carry out appropriate practical reasoning (Northcutt, 2007, p. 23). More importantly, the Darwin Awards illustrate people's intuitive sense of when it is not appropriate to expect an agent to take himself to have reason to act in some way. Some accidental deaths and disasters occur because agents fail to determine what they have objective normative reason to do, but are not a sign that the agents were not justified in taking themselves to have reason to act as they did. A man who died after he asked a friend to punch him to try to cure his hiccups was disqualified from the Darwin Awards because it would have been inappropriate to expect him to realize that a punch to the chest would kill him (Northcutt, 2007, p. 181). Accidental deaths involving children are also disqualified because children are agents-in-training (Northcutt, 2007, p. 4). These awards are one of many examples that demonstrate people's developed sense of what it is appropriate to expect human agents to realize they have reason to do.

Consider Bernard Williams' example: 'The agent believes that this stuff is gin, when it is in fact petrol. He wants a gin and tonic. Has he a reason to mix this stuff with tonic and drink it?' (Williams, 1981a, p. 102). Williams says, 'No', and many philosophers agree with him.[40] I say that if the agent's practical and epistemic circumstances mean that he would be justified in concluding both that there is gin in the bottle and that he has reason to drink it, he has justifiable reason to drink it.

---

[40] Williams' position is discussed in Chapter 7. Most of the philosophers discussed in this thesis agree with Williams on this point. See, also, Chapter 8.

Assume that an agent has an unproblematic desire for a gin and tonic and wants to mix the stuff in the bottle in front of him with tonic and drink it. If we take account of the agent's practical and epistemic circumstances to determine what the agent has justifiable reason to do, we need to find out whether the agent could realise that the stuff is petrol and understand the likely consequences of drinking it. The practical and epistemic circumstances that affect the agent's ability to come to learn relevant information will include his environment, personal situation, intellect, the nature of the belief, and what the agent can be motivated to do or learn. It should be apparent from this list that what an agent has justifiable reason to do will usually be affected by other justifiable reason relations that apply to him. As is often the case in philosophy, when Williams introduces this example, many important factors are left unexplained.

First, add in circumstances that would make it relatively easy for the agent to learn that there is petrol in the gin bottle. Assume that he has full use of his faculties. He is a mechanic standing in his friend's garage, with his friend next to him, looking at a half-full gin bottle sitting on a shelf. Next to the gin bottle is a rum bottle containing some blue liquid. In this situation, the mechanic has every reason to doubt that the stuff in the bottle is gin, and both his friend and his nose are available to check what the substance actually is. It is obviously inappropriate for him to conclude that the half-full gin bottle is a reason for him to mix the stuff in the bottle with tonic and drink it.

In contrast, assume that a student has gone home with severe hay fever; his nose is blocked, and his eyes are scratchy and blurry. The kitchen is a mess. A gin bottle stands open on the kitchen bench with a funnel next to it. He knows that his roommates tend to put stuff in any handy container, so the bottle could have been filled with something other than gin. But, the gin might be about to be used in cooking. There is no one home to ask, and he has no sense of smell. Is it appropriate for the student to conclude that the gin bottle gives him reason to drink from the bottle? Even though it would not be easy for him to check its contents, he has reason to doubt that the stuff really is gin. Past disasters may have left him strongly motivated to avoid risks when consuming stuff lying around the apartment. Although it is less clear than the previous case, he seems not to have justifiable reason to drink from the bottle.

Finally, consider a case where an agent could justifiably drink the stuff in the bottle. A blind agent with no sense of smell has received his grocery delivery from his reliable supplier. He ordered gin, and wants a gin and tonic. He feels for the correct bottle in the box. Unbeknownst to him, he has ended up with a gin bottle that a prankster at the bottling factory filled with petrol. Is it appropriate for him to take himself to have reason to drink the stuff in the bottle? He could telephone a chemist to come and check what is in the bottle, so it is logically possible for him to carry out the practical and epistemic processes necessary for him to learn that there is petrol in the bottle. But, it would be inappropriate – even paranoid – for the man to think that he had any reason to telephone someone to check the contents of the bottle. This agent has justifiable reason to mix the contents of the bottle with tonic and drink.

The practical and epistemic circumstances of the mechanic, the student and the blind man affect what they have justifiable reason to do. The mechanic has a justifiable reason not to mix the stuff in the bottle with tonic and drink it; it would be appropriate for him to check the contents of the gin bottle before doing anything with it. Ignorance of some state of affairs does not mean that an agent has no reason to take that state of affairs into account when he decides how to act, because agents can be culpable for their ignorance. In contrast, it is not appropriate to expect the blind man to check the contents of the gin bottle, and he has a justifiable reason to drink from it. What the student has justifiable reason to do is harder to determine. We need more information to learn whether he has a justifiable reason to drink from the bottle, and even with all the information available, we might not be able to state what he has reason to do.

One difficulty with trying to explain what it is appropriate for agents to take themselves to have reason to do is that we decide what it is appropriate for agents to take themselves to have reason to do by considering the reasons they have for investigating their reasons for action. As I mentioned above, agents have justifiable reasons for investigating what they have justifiable reason to do. So, whether an agent has a justifiable reason to act will be affected by whether he has justifiable reason to act in ways that will help him determine what he has justifiable reason to do. Recall the expanded formulation of justifiable reasons:

> *A* has justifiable reason to *φ* in *C* if and only if, were *A*, in *C*, to reason in a way that is possible and appropriate, she would be aware of *S*, and conclude that *S→(φ→E), E→V*, and *(φ→N)→¬(N>sigV)*.

An agent has justifiable reason to investigate whether he has justifiable reason to act in some way if and only if the probable disvalue of him checking what he has reason to do does not significantly outweigh the probable value of him checking up on his reasons. Given the mechanic's circumstances, the probable disvalue of him asking his friend about the contents of the bottle would not significantly outweigh the probable advantage of checking its contents. Given the blind, asnomic man's circumstances, the probable disvalue of him phoning a chemist to test the contents of the gin bottle value significantly outweighs the probable value of him learning the results of those tests. Hence, the blind, asnomic man has no justifiable reason to check the contents of the bottle, and, given his circumstances, he has reason to drink from it.

Agents' motivational systems can also affect what they have justifiable reason to do.[41] People's motivations are not always under their conscious control, and this lack of conscious control affects what agents can learn they have reason to do. If an agent cannot learn she has reason to act some way, it is not appropriate for her to take herself to have reason to act in that way. So, if it is in practice impossible for an agent to be motivated to carry out the actions or reasoning necessary for her to learn she has reason to act in some way, the agent does not have a justifiable reason to act in that way.

---

[41] I argue in Chapter 7 that this does not make my account of justifiable reasons an internalist account of reasons.

Motivation is sometimes treated as a goal-setter. Motivation in this sense is a pull to achieve a goal. When Williams writes about Owen Wingrave's lack of motivation to join the army, he refers to motivation in this goal-setting use of the term; Wingrave cannot be brought to set himself the goal of joining the army to uphold the honour of his family (Williams, 1981a, pp. 106-107). We usually expect this pull factor, end-setting motivation to be accessible by reason – even if someone thinks that the reasoning has failed to achieve the correct results. But, motivation also has a push factor. Someone can truly take themselves to be motivated to achieve a goal, but find it difficult to get motivated to do anything about it. This push factor motivation is harder to control through reason, or so I will argue; apart from anything else, scientists and other researchers are still learning about the things that affect it. One-marshmallow and two-marshmallow children provide a rough example of the difficulty of controlling push factor motivation (a rough example because of the age of the children) (Shoda, Mischel, & Peake, 1990). In the experiment, marshmallows are put in front of a child and the child is told that she can eat two marshmallows after fifteen minutes, but she will only get one if she does not wait. Even if all children set themselves the goal of waiting, not all children will wait. This may be part of what happens in some other cases of weakness of will. It is this sense of motivation as a push to act in certain ways that I am interested in here.

When people think of motivation in the pull factor, goal-setting way, they seem to think that when an agent is motivated, she usually consciously wants to act some way, and she is to some degree responsible for her motivations. But, motivation in the push-factor sense often seems to be outside of individuals' immediate control. Individuals can often train themselves in ways that affect motivation in the push-factor sense, but it takes knowledge, time, planning and conducive circumstances to do so. For example, alcoholism involves push-factor motivation, so an alcoholic might not be held responsible for her desire to drink. But alcoholics can learn that they are driven to drink to excess, and they are usually thought responsible for motivating themselves to control their drinking. There is an expectation that they can, and ought to, set themselves the goal of avoiding alcohol. If motivation in the pull-factor, goal-setting sense involves consciously desiring something, and people are responsible for directing their motivations, then agents' pull-factor motivational tendencies are unlikely to limit their practical and epistemic reasoning, and unlikely to affect what they have reason to do. But, social psychology experiments show that push-factor motivation is often affected by agents' circumstances without those agents being aware of the effect of their circumstances on their motivation. I give an example of this below. Thus, circumstances that affect a person's motivation may affect what that person is able to learn she has reason to do, and so affect what that person has justifiable reason to do.

Experiments on problem solving show strong relationships between features of motivation that cannot usually be controlled by agents and what agents are able to determine they have reason to do. For example, offering agents money or other rewards for solving problems can increase the time it takes for the agents to solve problems, or even prevent them from solving them (Amabile, Hennessey, & Grossman, 1986; Eva, 2010, p. 19; Glucksberg, 1962). Extrinsic rewards – 'extrinsic motivators' – seem to focus agents on completing the task

rather than on thinking of the best method for completing the task. Given that agents know the link between completion and the method for completion, and given that they want the reward, this effect on motivation seems to be unconscious, and so outside of agents' control. It seems that agents have capacities that they cannot exercise, or have difficulty exercising, because of the effect of circumstances on their motivational systems.

Experiments on the relationship between gender and empathy also seem to show that motivation does not always involve conscious, controllable desiring (Fine, 2010, pp. 20-21; Ickes, 2003; Ickes, Gesn, & Graham, 2000; Klein & Hodges, 2001). Ickes reports on a series of experiments measuring men's and women's ability to identify another person's thoughts and feelings (2003, pp. 119-151). The experiments were originally expected to confirm what 'everybody knows', that is, women are better at identifying other people's thoughts and feelings than men. To the experimenters' surprise, they initially found no difference between genders. Other research groups repeated the experiments and also found no significant difference between men's and women's ability to empathise. But, years later something changed. Three experiments that were only slightly different from the original sets of experiments produced results that the stereotypes predict. Suddenly women were significantly better than men at identifying what others were thinking and feeling.

The new experiments were different in one way: they asked participants to estimate the accuracy of their identification of the thoughts and feelings of the other person as well as asking them to identify the other person's thoughts and feelings (Ickes, 2003, pp. 125-131). This question somehow increased women's empathic accuracy. The researchers hypothesised that asking people how good they were at empathising reminded them of stereotypes that claim that women are better than men at understanding others. They found evidence from other studies that seemed to show exactly this kind of result (Ickes, 2003, pp. 130-132). It seems that such reminders somehow motivate women to try harder to identify the other person's thoughts and feelings. The researchers conclude:

> Although women, on average, do not appear to have more empathic *ability* than men, there is compelling evidence that women will display greater accuracy than men when their empathic *motivation* is engaged by situational clues that remind them that they, as women, are expected to excel at empathy-related tasks. (Ickes, 2003, p. 135, italics in original)

If the researchers are right to claim that 'situational clues … remind them that they, as women, are expected to excel at empathy-related tasks', this experimental result seems related to those that show that people tend to try harder when they believe they will be able to do something well, and tend not to try as hard when they think they are unlikely to succeed (Eccles & Wigfield, 2002; Ickes, 2003, p. 135; Wentzel & Wigfield, 1998).

I assume that agents in the experiments above set themselves what they would have thought to be the same goals; that is, in the motivation experiments, they all set themselves the goal of empathising to the best of their abilities. But beliefs about capabilities – and stereotypes – seem to affect the effort people put into things. Agents have some control over the beliefs that they form about themselves, whether that amounts to believing that, as women, they are

thought to be more capable of empathising, or to believing that, as women, they are thought to be less mathematically able. These agents cannot simply correct their beliefs about their capabilities or about stereotypes. (Interestingly, if they corrected their beliefs about men's and women's empathic abilities, they would probably have become *less* able to empathise.) First, forming justified beliefs about your abilities can be difficult, and agents sometimes need to decide how they have reason to act in situations where taking time out to research their capabilities would be inappropriate. Second, justified beliefs can be false. Even if an agent has the best possible understanding of her abilities, she may still limit her motivation by failing to understand her capabilities. Finally, this is relatively recent research, and many people do not know about it or understand its importance. The results of these experiments came as a surprise to the researchers and many others. Prior to these experiments it would arguably have been inappropriate to expect women to remind themselves that they were women before they tried to empathise with anyone. But it seems that if women remind themselves that they are women, they are better able to empathise. It seems unlikely that researchers in this area have finished learning all there is to know about the effect of push-factor motivation on actions.

The connection between this series of experiments and the claim that people's motivational tendencies sometimes unconsciously and uncontrollably affect their ability to empathise may seem tenuous, but another set of experiments reinforces this interpretation of the results. Kristi Klein and Sara Hodges carried out two different experiments on gender and empathy (2001). The first experiment replicated the results that Ickes and his colleagues found in the priming experiments (Klein & Hodges, 2001, pp. 721-725). Participants were asked to watch a video of a student discussing a difficult situation. Next the participants filled out one of two questionnaires. One questionnaire asked the students questions about their sympathy for the student. The second questionnaire was the psychological equivalent of a placebo pill. After filling out their questionnaires, the participants were shown the video again and asked to identify the student's thoughts and feelings at various points in the video. Sure enough, women who filled out the sympathy questionnaire before the empathy task showed better empathic ability than men in either group and than the women who did not fill out the sympathy questionnaire. Just as asking participants how likely it was that they had correctly identified someone else's thoughts and feelings seemed to motivate women to try harder to empathise, asking participants questions designed to elicit sympathy seemed to remind women that they were supposed to be able to understand others' thoughts and feelings and motivate them to try harder to empathise.

Klein and Hodges then carried out an experiment designed to test whether they could improve men's motivation to empathise (2001, pp. 725-727). The researchers ran the original experiment again, but before they gave the participants the sympathy questionnaire they told one group of participants that, depending on how accurate they were, they could receive up to two dollars every time they correctly identified the student's thoughts and feelings. Money is an external motivator. Men and women who were paid to correctly identify the student's thoughts and feelings were significantly better at empathising than men and women who

were not paid, and the gender-related difference in empathy became insignificant (Klein & Hodges, 2001, p. 727). Klein and Hodges write:

> In sum, motivation seems to be a key component in the process of empathising with another person. The gender difference that we found in our studies seems to be adequately explained by differential motivation between men and women under different circumstances. (Klein & Hodges, 2001, pp. 728-729)

Just as reminding women that, as women, they are good empathisers triggers the internal motivational tendency people have to try hard when they believe they are likely to succeed, offering men and women money to empathise serves as an external motivator, spurring them to try harder to empathise.

People's ability to empathise with others will affect what they have justifiable reason to do, because other people's mental states can serve as reasons for our actions. For example, if someone laughs in response to a story of woe, whether the laugh appears sympathetic or belittling will affect what the storyteller has reason to do. The research described above suggests that minor changes in circumstances affect people's motivations in ways that affect the accuracy with which they identify others' mental states. This suggests that motivation affects what it is appropriate for agents to take themselves to have reason to do. The degree of conscious control that people have over such motivational nuances is unclear. In the above experiments, the participants were presumably unaware of the ways in which their circumstances affected their motivation, and if they were unaware of the way the circumstances affected their behaviour, it is unlikely that they could have altered their behaviour. Even if they knew that their behaviour was likely to be affected by their circumstances, they may not have been able to control the effect of their circumstances on their behaviour.

The experiments on problem solving and empathy described above seem to show that agents' circumstances sometimes affect their motivation in ways that they will be unaware of, and, hence, in ways that it is inappropriate to expect them to control. If this is what they show, then motivation can affect what it is possible for agents to learn they have reason to do, and so motivational abilities will affect what agents have justifiable reason to do.

My claim that agents' practical and epistemic circumstances affect what they have reason to do is not a claim that agents do not have justifiable reasons to do things that it would be psychologically difficult for them to learn they have reason to do. I am not claiming, for example, that agents should never put more effort into empathising with others than they currently do. Even when it is difficult for an agent to learn something about the world, values, or his motivations, the agent may still have reason to learn about that state of affairs, and that state of affairs may help determine what he has reason to do. It is not the mental or physical effort needed for the agent to learn that something that affects what an agent has justifiable reason to do, but whether it is possible and appropriate for the agent to seek to learn that information. Whether an agent has a justifiable reason to do something that it is difficult for him to learn he has reason to do will depend on such things as the importance of the action and potential consequences of an error. It may, for example, be psychologically difficult for a

man who was abused as a boy to learn that he has reason to consider the effects of abuse on others when deciding whether he has reason to abuse them. Nevertheless, the seriousness of the potential consequences means that the agent is likely to have justifiable reason to make the effort to think such issues through.

Just as I do not claim that agents have no justifiable reason to do things that it would be psychologically difficult for them to learn they have reason to do, I don't claim that agents don't have reason to do things that it would be time-consuming for them to learn they have reason to do. Again, the effort that it is possible and appropriate for an agent to expend to learn what he has reason to do will be affected by the disvalue associated with him using his time in that way and the disvalue associated with him failing to get it right. If an agent puts a lot of effort into establishing that the person who collapsed next to him is not faking, the potential harm caused by the failure to help immediately is great. If the agent helps the collapsed person straight away, the potential harm caused by helping someone who does not need help is likely to be much lower – at least it would be appropriate for the agent to think that if he were to spend the time weighing the options.

4.4 Relationships: justifiable reasons and other forms of reasons

So far in this thesis I have distinguished between justifiable reasons and three established categories of reasons: motivating reasons; normative reasons; and objective normative reasons. The relationship I take to hold between justifiable reasons and normative reasons is clear: I take it that justifiable reasons are normative reasons. Arguments against this are considered in the next chapter. The relationship between justifiable reasons and motivating and objective normative reasons may be unclear. Although there are differences between justifiable reasons and motivating and objective normative reasons, I take it that one and the same reason for acting may be motivating, justifiable and an objective normative reason for an agent to act.

As implied by their normativity, justifiable reasons are not the same as motivating reasons. The statement, 'R is a justifiable reason for A to φ, in C', does not mean 'R leads, or would lead, A to believe that some reason justifies her φ-ing'. This would not be a statement about normative reasons, because an agent could take herself to be justified in acting some way after carelessly flawed reasoning. When an agent takes herself to have reason to act in some way after carelessly flawed reasoning, she has a motivating reason, but not a justifiable reason. Reasons and reason relations may also be justifiable without motivating, such as when an agent would realise that she has reason to act in some way were she to think about her situation in a way that is possible and appropriate, but she never engages in the appropriate reasoning.

Justifiable reasons and objective normative reasons also differ. What an agent has justifiable reason to do is determined by what the agent would have reason to do were she to reason in a way that would be possible and appropriate given her circumstances. What an agent has

objective reason to do is determined by ideals that are in some ways independent of the agent's psychological or contextual limitations. If someone next to you fakes a collapse and you cannot learn that the person is faking, you usually have justifiable reason to try to help, even when you would not have objective reason to do so. If someone has secretly tipped a little petrol into a bottle of gin at a party, you may have justifiable reason to drink it, even though you would not have objective normative reason to do so. Objective normative reason relations are derived from a highly idealised perspective on agency or an ideal understanding of the world. In contrast, statements about justifiable reasons assume that a human agent functions as ideally as that agent could realistically be expected to function in her circumstances.

Although they are distinct concepts, an agent's objective normative, justifiable and motivating reasons can coincide. When an agent acts on the basis of a justifiable reason, the agent's motivating and justifiable reasons and reason relations will be the same. A justifiable reason statement can also correspond to an objective normative reason statement when the information and reasoning available to the agent corresponds with whatever form and degree of idealisation is thought necessary for determining what the agent has objective normative reason to do. Consider Bernard Williams' account of objective normative reasons. According to Williams, if there is petrol in the gin bottle, an agent who wants a gin and tonic has a normative reason not to pour the stuff into a glass, mix it with tonic, and drink it (1981a).[42] I argued above that there are circumstances, such as that of the mechanic, where the agent also has a justifiable reason not to pour petrol into a glass and drink it. In such a case, the agent's justifiable and objective normative reasons coincide. If, in contrast, an agent is taken to have objective normative reason to act in whatever way leads to the best outcome, there are likely to be circumstances where agents have justifiable reasons for acting in ways that lead to the best outcome. Similarly, what an agent has justifiable reason to do could also coincide with what the agent has reason to do on a constructivist account of objective normative reasons. So, when an agent's knowledge of himself and the world is ideal in whatever ways are required by the relevant account of objective normative reasons, the justifiable and objective normative reason statements that apply to the agent will be the same.

As explained in §3.3, the reasons that generate reasons relations are, metaphysically speaking, the same kind of thing whether they are the basis for motivating, justifiable, or objective normative reason relations. I assume that reason-explanans identified by '*STATE*', are states of affairs. Given that an agent can act for a reason that is simultaneously a motivating, justifiable and objective normative reason, there would be a tension in my theory if I held that what reasons are differs for the different reason statements. What differs for motivating, justifiable, and objective normative reason-explanans are the other aspects of the reason relation that pick out those particular states of affairs and make them reason-explanans rather than mere states of the world. As I argue in Chapter 8, objective normative reason relations identify actions that will yield a result that has value, on a very broad interpretation of the

---

[42] In Chapter 7, I argue that Williams' account of normative reasons is an account of objective normative reasons.

word 'value'. Which state of affairs serves as *STATE* in an objective normative reason relation depends on a suitably idealised account of *END, VAL,* and *SIG*. Justifiable reason relations explain what real human agents have pro tanto normative reason to do. Which state of affairs serves as *STATE* for a justifiable reason relation depends on what it is appropriate for the agent to take as *END, VAL*, and *SIG*.

When there are different concepts of what seem to be the same kind of thing, philosophers tend to want to know which concept, if any, is foundational. Hence, the concepts of subjective normative reasons considered in Chapter 6 are taken to be beliefs about objective normative reasons, and objective normative reasons are usually treated as foundational.[43] In contrast, I take it that justifiable and objective normative reasons and reason relations are distinct concepts that serve different functions, and questions about which are foundational are misguided.

Justifiable reasons are not a subset of objective normative reasons, that is, agents do not have justifiable reason to do whatever they can determine they have reason to do from the set of objective normative reasons.[44] As in the above examples, an agent can have justifiable reasons for acting in ways that conflict with what he has objective normative reason to do. Thus, the blind man can have a justifiable reason to drink petrol, even though he has no objective normative reason to do so. Similarly, objective normative reasons are not a subset of justifiable reasons. The idealisation involved in determining what an agent has an objective normative reason to do means that an agent can have an objective normative reason to act in ways in which he has no justifiable reason to act. The example of the blind man is also an instance where this occurs.

Justifiable reasons and reason relations are also not beliefs about what an agent has objective normative reason to do.[45] They are not even justified or justifiable beliefs about what an agent has objective normative reason to do. First, the existence of justifiable reasons and reason relations is independent of anyone's beliefs about what they have reason to do. To describe justifiable reasons as 'justifiable beliefs about objective reasons' invites the mistaken conclusion that it is the beliefs that determine what someone has a justifiable reason to do. To take account of this, justifiable reasons would need to be described as beliefs that agents would be justified in forming about what they would have objective normative reason to do. However, even this attempt to explain justifiable reasons in terms of objective normative reasons fails, because an agent can have justifiable reason to act in some way when it would be inappropriate for her to believe that she has an objective normative reason to act in that way. In situations of uncertainty agents may not be justified in forming any belief about what

---

[43] I argue in Chapter 6 that this approach is mistaken if subjective normative reasons are intended to explain real agents' reasons for action.

[44] Hence, I take it that justifiable reasons do not serve the same kind of function as the reasons that pass through Jonathan Dancy's epistemic filter. This distinction between justifiable reasons and objective normative reasons is discussed further in Chapter 8.

[45] This is discussed further in §6.3 and Chapter 8.

they would have objective reason to do. Someone collapses next to you and stops breathing. Your rusty knowledge of first aid may mean that you have only a vague recollection of the correct resuscitation technique. So, you would not be justified in believing that you have an objective normative reason to carry out any particular resuscitation procedure. You may not even be justified in believing you have an objective normative reason to attempt to resuscitate the stranger. You are ignorant of too many relevant truths, and there is little time for you to carry out any form of ideal deliberation. However, that same rusty knowledge may justify you attempting to resuscitate the stranger. So, you may have a justifiable reason to attempt resuscitation, even though it would not be appropriate for you to believe that you have an objective reason to carry out some resuscitation procedure.

Errol Lord gives an account of reason-explanans that is very similar to my account of justifiable reason-explanans, but seems to try to make such reasons a form of objective normative reason. Lord develops what he calls, following Mark Schroeder, a 'factoring account' of objective normative reasons:

> Agent A *has reason R* to (intend to) *p* iff (1) *R* fits the Practical Model and (2) *A* is in a position to know *R*. (Lord, 2010, p. 287)

Where a proposition fits the Practical Model if that proposition is a practical reason (Lord, 2010, p. 287, fn 9). Lord states that he is 'mostly neutral as to what it takes for some proposition to fit the Practical Model', but that he assumes 'that practical reasons are … factive' (2010, p. 287, fn 9). There are commonalities between my account of reasons and Lord's factoring account. I take reasons to be states of affairs rather than propositions, but this difference is of little importance here. For a state of affairs to be a justifiable reason explanans, that is, for it to correspond to Lord's *R*, it must be appropriate for the agent to be aware of that state of affairs. If it is appropriate for an agent to be aware of a state of affairs, then it seems plausible that it would also be appropriate for the agent to form a belief about that state of affairs, which is a significant step towards claiming that the agent is in a position to know about the state of affairs. Lord makes almost no claims about the nature of the reason relation that would be specified by what he calls the 'Practical Model'. So, my account of reason relations could serve as the relevant Practical Model, which would make Lord's factoring reasons and my justifiable reasons even more similar. However, Lord's lack of specifications for the Practical Model for his account of reasons undermines his claims about factoring reasons.

Like me, Lord claims that his account of factoring reasons is an account of reasons with a close relationship to practical rationality (2010, pp. 293-294).[46] However, one of Lord's prime concerns is establishing the relationship between his factoring reasons and Mark Schroeder's objective normative reasons (Schroeder, 2008b, 2008c). So, although Lord says almost nothing about the Practical Model, his arguments tend to assume that he can insert Schroeder's Practical Model, or some other account of objective normative reason relations,

---

[46] Lord argues that factoring reasons 'help clarify and explain substantive rationality' (2010, p. 293, italics removed). However, Lord's concept of 'substantive rationality' is atypical. Differing concepts of rationality are discussed in §9.1. Lord's use of the term is clarified in a footnote in §9.1.

into his factoring account (Lord, 2010). If Lord wishes to keep the link he suggests exists between his factoring reasons and rationality, he needs to exclude Practical Models of objective normative reasons from his factoring account of reasons. As I argue in Chapter 8, accounts of objective normative reasons are concerned not only that reason-explanans, such as *R*, are states of affairs or true propositions, but also that reason relations properly represent values, where, as always, the word 'value' is to be interpreted broadly. If Lord's Practical Model satisfies the requirements on means-end relationships for objective normative reasons represented in my account of reasons by *END*, and the requirements on value relationships for objective normative reasons represented by *VAL* and *SIG*, his factoring account will be objective in the relevant sense, but it will become detached from rationality. If Lord adopts the Practical Model I set for justifiable reasons, then he must accept that his concept of objective normative reasons is very different from that at issue for Schroeder, whose work is Lord's principle focus.

It would be a little quick to conclude from these arguments that justifiable reasons cannot be explained in terms of objective normative reasons, but I have yet to find any way in which this can be done using current conceptions of objective normative reasons.

4.5 Conclusion

Justifiable reasons are normative reasons that real human agents can determine they have reason to act on. Justifiable reasons have a significant advantage over motivating and objective normative reasons; they represent the form of reasons that real agents have reason to act on, and figure in accounts of appropriate self-regulation, praise, blame and rationality. An agent has a justifiable reason to act in some way when she would have reason to act in that way were she to reason about her circumstances using methods that it would be both possible and appropriate for her to use. Justifiable reason-explanans are states of affairs that agents ought to be disposed to take to give them reason to act in a certain way. What makes justifiable reasons normative is their justifiability.

The notion of 'possible and appropriate reasoning' on which my concept of justifiable reasons relies is vague. At times it is unclear whether it is possible and appropriate for someone to decide she has reason to act in some way. The situation is further complicated by the dependence of justifiable reasons on other justifiable reasons. Determining what someone has justifiable reason to do sometimes requires determining whether the agent has justifiable reason to carry out investigations to determine whether she has justifiable reason to act in some way. In spite of issues with complexity and vagueness, we do have a common-sense understanding of what it is possible and appropriate for agents to determine they have reason to do, and any vagueness corresponds to real uncertainties about what agents have good reason to do.

An agent can have a justifiable reason to act in the same way that she has motivating and objective normative reason to act. However, justifiable reasons and reason relations cannot be reduced to motivating or objective normative reasons or reason relations.

5. Objections to Justifiable Reasons

5.0 Introduction

In this chapter I consider five possible objections to justifiable reasons. Only one of these arguments is directed against my account of justifiable reasons, but the others are directed at accounts of reasons that in some way resemble my account. I consider three objections that claim that justifiable reason relations and reason explanans are not normative, an objection that claims that my account of justifiable reasons is the same as Dancy's account of normative reasons, and an objection that claims that my account of justifiable reasons is not a viable account of practical reasons because it excludes reasons that real agents would be disposed to treat as pro tanto reasons if they reasoned in a way that was possible and appropriate. I argue that my account is not vulnerable to these objections.

Peter Railton claims that first-personal accounts of practical reason cannot provide a complete explanation of practical reason because first-personal accounts cannot include all of the information needed for a full account of practical reasons. He describes first-personal approaches to reasons as based on considerations about 'whether the agent makes appropriate use of information and motives actually available to her in the situation' (Railton, 2009, p. 82). Railton's description of first-personal accounts of practical reason resembles my account of justifiable reasons, so in §5.1 I show that Railton's objection to first-personal accounts of practical reason does not apply to my account of justifiable reasons because my account of justifiable reasons is not a first-personal account of reasons.

The next two objections to justifiable reasons arise from arguments that only objective normative reasons are genuine normative reasons. The objections discussed in §5.2 stem from arguments that claim that only reasons that are appropriately grounded in reality are normative. According to the first version of this objection, objective normative reason-explanans are reasons because they are normally states of affairs, but justifiable reason-explanans must be beliefs. So, in §5.2.1, I defend my claim that justifiable reasons are states of affairs. According to the second version of this objection, described in §5.2.2, justifiable reason-explanans may be states of affairs, but those states of affairs are not picked out as reasons through a process that is appropriately responsive to reality. I argue that this argument does not show that justifiable reasons are not normative.

Work by Jonathan Dancy leads to two radically different objections to my account of justifiable reasons. Unlike most of those who develop accounts of objective normative reasons, Dancy argues that only states of affairs that are accessible to agents can be normative reasons. Thus, it might be claimed that my account of justifiable reasons is an attempt to supersede Dancy's account of normative reasons, and should be treated as such. So, in §5.3.1, I show that there are important differences between my account of justifiable reasons and Dancy's account of objective normative reasons. Finally, in §5.3.2, I consider Dancy's

argument that there cannot be two forms of normative reason, so given that we know that there are objective reasons, there cannot be justifiable reasons. I argue that there can be, and are, at least two forms of normative reason, but that justifiable reasons are the only form of normative reason that agents can act on.

5.1 Railton's objection to first-personal accounts of practical reason

Peter Railton's arguments about the importance of practical competence and fluid agency in practical reason might be thought to raise an objection to my account of justifiable reasons (2009). One conclusion Railton draws from his arguments is that it would not be possible to give a good account of practical reason by considering agents' situations solely from a first-personal perspective. My account of justifiable reasons might be thought to attempt to do just that, but I argue my approach is, in fact, close to the approach that Railton recommends, combining both a first-personal and third-personal approach to understanding reasons for action.

Railton distinguishes between theories of practical reason that take what he calls an 'internal', 'first-personal approach' and those that take an 'external', 'third-personal standpoint' (2009, p. 82, italics removed). Railton's 'first-personal approaches' to understanding practical reason consider 'whether the agent makes appropriate use of information and motives actually available to her in the situation' (2009, p. 82). This description of first-personal approaches to practical reason seems to resemble the approach I take in my development of an account of justifiable reasons. In contrast, the 'third-personal standpoint of explanation' involves explaining reasons 'from a standpoint "external" to the agent's own perspective' (Railton, 2009, p. 82). Railton describes this as an 'objective' approach to understanding practical reason, and at first glance it seems to resemble the approach taken by those who argue that objective normative reasons are the only form of normative reasons (Railton, 2009, p. 83). Railton gives a comprehensive and interesting argument for this distinction between first- and third-personal approaches to practical reason.

Railton uses recent work from psychology and cognitive science to argue that many of our intentional actions arise from processes that are autonomous, but do not involve conscious deliberation.[47] These processes affect both our identification of states of affairs as reasons and our decision to carry out actions. Agents can correctly identify states of affairs that are normative reasons without conscious awareness that they have taken those states of affairs to be reasons. Thus, a nurse may respond to subtle clues that something is going wrong with a patient without being aware which particular state of the patient triggered her action. At times, agents will not be able to pick out the precise state of affairs that they took to be a reason even after reflection. When asked, the nurse may say she could 'just tell' that the patient needed help and claim that she could not identify the specific collection of features that

---

[47] John Broome also gives a short argument for something like this view, although his argument relies on intuitions rather than science (2007a, pp. 355-356)

served as a reason to act. In situations where agents are unable to identify the states of affairs that they took to be reasons, outside observers with suitable training can sometimes identify the reasons to which the agents responded. People who are experts at detecting and responding to reasons are sometimes studied by expert analysts so that their skills in identifying reasons for action can be written down and taught to others (Johnson & Johnson, 1987; Wright & Ayton, 1987).[48] As well as unconsciously identifying that states of affairs are reasons, agents can decide what those states of affairs give them reason to do without an awareness of making a decision. Important aspects of an agent's response may also be outside the agent's ken. Again, a suitably trained observer may be better at identifying and explaining what the agent decided to do than the agent herself. Many fields have examples of agents exhibiting such skills of practical reason, and Railton lists some of the skills involved (2009, pp. 108-110).

Railton suggests that first-personal approaches to understanding practical reason are unable to capture everything involved in practical reasoning because agents will often be unaware of their identification of elements of the world as reasons for action and their so-called 'decision' about what they have reason to do (2009, p. 113). Railton also claims that third-personal approaches have limited scope for explaining practical reasons, but this is unlikely to concern any objective normative reasons theorists who are not concerned about whether their theories reflect real people's reasons for action. Unlike objective normative reasons theorists, I aim to give an account of normative reasons that people can respond to, so if Railton's criticisms of first-personal approaches apply to justifiable reasons, his criticisms affect my theory.

In spite of surface appearances, my account of justifiable reasons is not a first-personal account of reasons or a first-personal perspective on reasons; instead, it combines a first-personal and third-personal approach. Justifiable reasons are states of affairs that it is appropriate for agents to take to be reasons. This does not mean that agents need to be consciously aware of those reasons or able to identify them after reflection. As mentioned in §3.1, I accept Railton's arguments that normative reason-explanans and normative reason relations can operate without an agents' consciously deliberating about those reasons or even being able to express them after reflection. My account of reasons takes the perspective of the agent in the sense that the states and abilities of agents will limit what states of affairs serve as normative reasons. But, my account of reasons is not 'first-personal' in the sense that explanations of reasons must be given as though they come from the perspective of the agent. Not even accounts of motivating reasons need to be 'first personal accounts' of reasons in this sense. As Railton suggests, an agent may only come to understand what justified her acting in a certain way after someone else adds an explanation from a third-personal standpoint to her first-personal understanding of her situation (2009, p. 113). While my definition of justifiable reasons means that agents will always be able to determine what they have justifiable reason to do, it is compatible with this that at times agents will not know that they know what they have justifiable reason to do.

---

[48] Numerous examples of knowledge elicitation techniques, and analyses and uses of that data, are available in literature on the development of expert systems.

5.2 Justifiable reasons are not appropriately connected to reality

Objective normative reasons are sometimes said to be the only genuine form of normative reason for action, and this claim might be used to argue that justifiable reasons are not normative. Robert Audi writes, '[normative] reasons are objective', without feeling the need to argue for the claim (2001, p. 119). Jay Wallace writes that normative reasons are objective, and so-called 'subjective reasons' are not really reasons at all (2003a). Jonathan Dancy argues that normative reasons are objective inasmuch as they 'are, in a sense, independent of our [beliefs and] desires' (2000, p. 49). And, Michael Smith replies to Christine Swanton's criticisms of his position by arguing that on his account, 'normative reasons turn out to be thoroughly objective', implying that anything other than 'thorough objectivity' would be a barrier to normativity (Smith, 1996, p. 161). [49] Even Michael Smith, who argues that normative reasons result from coherent reasoning rather than ideal consequences, insists that the search for coherence begins with facts. Claims that normative reasons must be objective usually amount to claims that normative reasons are so-called 'real' reasons, or that normative reason relations are based on reality.

I claim that, like objective reasons, justifiable reasons are states of affairs, and that justifiable reason relations are just as appropriately tied to the actual state of the world as objective normative reason relations. There is, however, an obvious objection to my claims. It might be argued that justifiable reasons cannot be states of affairs because even when people have an understanding of the world that is appropriate given their circumstances, they can have false beliefs about states of affairs. In §5.2.1, I argue that this is not a successful objection to my claim that justifiable reasons are states of affairs. It might also be argued that even if justifiable reasons are states of affairs, the limited human reasoning that picks out those states of affairs as justifiable reasons and determines what agents have justifiable reason to do cannot tie justifiable reason relations to the world in a way that makes such reasons normative. States of affairs only become justifiable reasons when they are appropriately connected to agents' psychological states. An agent who is psychologically incapable of detecting that on a highly idealised understanding of her situation she has reason to carry out some action does not have justifiable reason to act in that way. This tie to individuals' psychological states means that agents may have justifiable reasons to act in ways that conflict with the actions that they have objective normative reasons to carry out. In §5.2.2, I show that this objection does not show that justifiable reason relations are not normative.

5.2.1 Normative reasons are states of affairs

Objective normative reasons are usually said to be either states of affairs or propositions, and the normativity of objective reason relations is sometimes taken to be derived from their grounding in the actual state of the world. Thus, normative reasons are sometimes described

---

[49] Michael Smith's position is discussed in Chapter 7.

as being 'real reasons', or as being 'based on reality'. When normative reasons are described in this way, they are contrasted with what are called 'subjective reasons'. Thus, Jonathan Dancy describes objective reasons as 'grounded in … features of the situations that we face' and subjective reasons as grounded in 'the world as we take it (or defensibly take it) to be' (2009, p. 97). Similarly, Jay Wallace argues that objective reasons are 'genuine normative reasons', while subjective reasons are 'beliefs of agents about what they have reason to do' and, so, 'not really reasons at all' (2003a). This claim that only objective reasons can be normative is the foundation of Wallace's argument against the existence of subjective normative reasons presented in §6.2.2. Although justifiable reasons do not correspond to any of the standard analyses of subjective reasons, as I explain below, this argument could be generalised to refer to justifiable reasons.

The claim that justifiable reasons are states of affairs is likely to be met with the same objection as the claim that motivating reasons are states of affairs, namely, the claim that such reasons can be false and, so, must be beliefs.[50] When an agent acts for a reason, the agent's motivating reason is whatever it is that the agent took to provide a reason for acting. With justifiable reasons, as with all forms of normative reasons, the reason the agent has for acting is whatever the agent ought to be disposed to take as an indication that there is reason to carry out a certain action.[51] However, justifiable reasons differ from other forms of reason-explanans because what the agent ought to be disposed to take to be a reason to act in some way is limited to what it is possible and appropriate for her to take to be a reason to act in that way. As with motivating reasons, agents' psychological states will affect what agents have justifiable reason to do, because agents' psychological states affect what it is possible and appropriate for them to take to be reasons. Something may not be a justifiable reason for an agent even though it is an objective normative reason; hence the petrol in the gin bottle is not a reason for the blind, asnomic man to not drink from the gin bottle. Something may be a justifiable reason for an agent even though it is not an objective reason. Even if someone faking a collapse next to you is not an objective reason for you to help him, it would normally be a justifiable reason for you to do so. Thus, just like motivating reason-explanans, justifiable reason-explanans can conflict with the reasons an agent would have if she had a highly idealised understanding of the world.

In the sense of the term that is relevant here, justifiable reason-explanans are states of affairs that it is possible and appropriate for agents to take as indications that they have justifiable reasons to act in some way. I am interested in the justifiable reason-explanans picked out by *STATE*, or '*S*', in the formula:

> *A* has justifiable reason to φ in *C* if and only if, were *A*, in *C*, to reason in a way that is possible and appropriate, she would be aware of *S*, and hold that $S \rightarrow (\varphi \rightarrow E)$, $E \rightarrow V$, and $(\varphi \rightarrow N) \rightarrow \neg (N > sigV)$.

---

[50] This argument against the claim that motivating reasons are states of affairs was discussed in §3.3.

[51] I argue for this claim in Chapters 2, 3, and 4.

I take '*S*' in this formulation of reasons to be a state of affairs, not the agent's beliefs. Consider an example from Cullity and Gaut (1997, pp. 1-2). Your nature and circumstances include that you are the kind of person who would read a philosophical thesis, that you have had pretty standard life experiences, and that you are ill. Your carefully chosen, well-qualified, and previously reliable doctor tells you to take a certain medicine. '*S*' is the state of affairs where your doctor tells you to take the medicine. '*φ*' is the action of taking the medicine. '*E*' is the effect that it is appropriate for you to expect will follow from the action, in this case improved health. '*V*' is the positive value of the action that it is appropriate for you to expect, in this case the expected value of improved health. '*N*' includes any negative side effects that it is appropriate for you to expect to experience if you take the medicine. Given your circumstances, it is both possible and appropriate for you to conclude that you have reason to take the medicine. You have a pro tanto justifiable reason to take that medicine, and it would be inappropriate for you to conclude that you have no reason to take it.

Unfortunately, your doctor got it wrong; he gave you a prescription for the wrong medicine. Those who argue that only objective normative reasons are states of affairs may claim that this means that the justifiable reason for you to take the medicine must be a psychological state of yours; perhaps your justifiable reason is your belief that you have reason to take the medicine or your belief that your doctor gave you good advice. However, given my definition of what it is for an agent to have justifiable reason to act in some way, there seem to be no grounds for saying that the reason-explanans for you to take the medicine is your belief. The justifiable reason-explanans for you to take the medicine is your doctor's act of telling you to take it, which is a state of affairs even when your doctor is wrong. Assume that you take the medicine. If you insist that your justifiable and motivating reason for swallowing it was your justified false belief that the medicine would help you, you are using 'reason' to refer to a different component of the reason relation than the one that is at issue here. Rather than using 'justifiable and motivating reason' to refer to *STATE*, you are using it to refer to *END*. This same move from using 'reason' to refer to *END* rather than *STATE* happens in other examples, hence, in part the importance of distinguishing between the various elements of the reason relation.

Lord, whose position was outlined in §4.4, gives a similar argument in support of the claim that the agent's reason is not a belief (2010, pp. 286-289). Lord claims that reasons are propositions rather than states of affairs, but this does not affect the key arguments. He argues that given that the agent wants to drink gin and tonic, the agent's reasons for believing that there is gin in the glass of petrol and tonic are also the agent's reasons for drinking the petrol and tonic. My account of reasons differs from Lord's, but the claim that states of affairs that are justifiable reason-explanans for an agent to act in some way will also be reasons for the agent to form a belief about that state of affairs seems correct. I hold that the states of affairs that give the agent reason to believe that there is gin in the glass of petrol and tonic also give the agent reason to drink the petrol and tonic when it is appropriate for the agent to take it that the relevant *END, VAL*, and *SIG* obtain; in this case, when it is appropriate for the agent to take it that: (*END*) if there is gin and tonic in the glass, drinking from the glass will satisfy his desire to drink G&T; (*VAL*) satisfying his desire to drink gin and tonic has positive value;

and (*SIG*) satisfying his desire to drink G&T does not have disvalue significant enough to outweigh the value of satisfying his desire to drink G&T.

Jeremy Fantl and Matthew McGrath develop an argument that could be used as an objection to my claim that states of affairs that would justify agents having false beliefs can be justifiable reason-explanans (2009, pp. 103-104). They develop an argument about motivating reasons, normative reasons, and what they call 'justifying reasons', reasons that would justify an agent's acting in a particular way (or believing a certain proposition) (Fantl & McGrath, 2009, pp. 103-105, 134-139). However, their arguments about justifying reasons appear to show that my claims about the contents of *STATE* must be incorrect (Fantl & McGrath, 2009, pp. 103-105, 134-139).

The simplest way to explain Fantl and McGrath's argument is through the use of an example. Gina in Gisborne and Petra in Paihia each go up to their usual bartender in their usual bar and ask for a gin and tonic. Each bartender picks up a gin bottle half-filled with clear liquid, mixes the contents of the bottle with tonic and passes it over. Gina and Petra both justifiably form the belief, 'I am holding a G&T'. Each takes herself to have reason to drink. Unfortunately, the gin bottle in Paihia contains petrol. Fantl and McGrath claim that the agents' motivating reasons are the contents of their beliefs. They point out that as the agents move to take a sip, Gina and Petra both have mental states with the content 'I am holding a G&T'; they do not have mental states with the content 'there is clear liquid in my glass that was poured from a gin bottle by my usual bartender'. Fantl and McGrath argue that it must be the content of the agents' beliefs when they act that motivate their action, not the various facts that led them to form those beliefs. They write: 'Anything further back in the justificatory chain wasn't the content of an operative belief' (Fantl & McGrath, 2009, p. 103). They claim that the content of the operative belief that Gina and Petra had when they drank was 'There is G&T in my glass'. So, Gina's and Petra's motivating reasons were 'My glass contains G&T', not 'My glass contains clear liquid'.

The connection between these arguments about motivating reasons and my claim that justifiable reason-explanans are states of affairs may seem obscure at this point. But, Fantl and McGrath, like Dancy and I, take it that the same things serve as reasons, whether they be motivating, justifying, justifiable or normative reasons. And in the example above, Gina and Petra act for a motivating reason-explanans that is also a justifiable reason-explanans. So, if Fantl and McGrath are correct, the agents' justifiable reason-explanans must be 'My glass contains G&T', not 'My glass contains clear liquid'. But if Petra's justifiable reason-explanans must be 'My glass contains G&T', and her glass contains P&T, Petra's justifiable reason-explanans looks a lot like a false belief.

Petra's position does not cause problems for Fantl and McGrath, because they adopt Dancy's argument that explanations of motivating reasons can be non-factive, which I discuss in §3.3, and integrate it into their account of justifying reasons (Dancy, 2000, pp. 121-137; Fantl & McGrath, 2009, pp. 103-105, 134-139). So, they claim that Petra's justifying reason for drinking the P&T would have been the proposition '[my glass contained G&T], as I thought

at the time' (Fantl & McGrath, 2009, p. 104). However, as mentioned in §3.3, it is debatable whether Dancy's non-factive explanation response succeeds, so I am less willing to accept this solution. Hence my move, as Fantl and McGrath put it, 'further back in the justificatory chain'.

Fantl and McGrath assume that agents' motivating reasons are reflected by what the agents would report to be the contents of their beliefs when they acted, but there are grounds for thinking that agents' motivating reasons are more complex than Fantl and McGrath's argument suggests. Agents' reports of their reasons for acting can be misguided. Gina and Petra might both insist that what motivated them was the thought that 'There is G&T in my glass' and not the thought that 'There is clear liquid that came through a mechanism reliably known to provide gin' without that entailing that the first was their motivating reason for drinking, not the second. Whatever Gina and Petra report about their reasons for acting as they did, Gina and Petra will have believed both claims. Moving up the justificatory chain may give a more accurate rather than less accurate account of the agents' motivating reasons, no matter what they would report their reasons to be.

There is a relationship between justifiable reasons and the potential contents of beliefs, but not the one that Fantl and McGrath suggest. In the usual bartender, usual bar, usual gin bottle, with usual gin-like appearance scenario, an agent would be justified in believing that there was gin and tonic in her glass. If she non-problematically wants a G&T, then she would also have justifiable reason to drink from the glass.[52] If she forms a belief that she has G&T in her glass, what justifies her belief is the situation in which she finds herself, including the clear liquid in her glass. Similarly, if she forms an intention to drink, what justifies her action is the situation in which she finds herself, including the clear liquid in her glass. In both the belief and the action situation, the clear liquid is one possible justifiable reason-explanans for believing or acting.

Cases such as your bartender pouring you a drink or your doctor prescribing you a medicine are reasonably straightforward. However, people sometimes treat perceptions as reasons when it is less clear that those perceptions report states of affairs. Some cases of illusions are easy to deal with while treating the motivating or justifiable reason-explanans as a state of affairs. If you want a straight stick for walking, and you see a stick poking half out of a clear, calm lake, the stick's looking bent where it enters the water is not a reason for you to decide that it will not suit your purpose. Most people over a certain age recognise the illusion of a bend created by differences in the diffraction of light by air and water. The stick looks bent, but that is an entirely appropriate state of affairs. The effect of sticks that look bent and other illusions on justifiable reasons are easier to deal with than delusions. Consider an alcoholic in withdrawal who is suffering from delirium tremens (colloquially, 'DTs') and who perceives fantastic, wicked creatures surrounding him.[53] In cases such as this, it is unclear what it is possible or appropriate for the agent to take himself to have reason to do. In extreme cases the

---

[52] This is, in part, what Lord claims (2010).

[53] Wonderful descriptions of people suffering from such delusions are given in (Brierre de Boismont, 1859).

person's ability to reason and consider his situation is so radically undermined that it is unclear whether he should still be considered an agent. However, it would be implausible to claim that deluded people are never agents, and cases of deluded agents appear to undermine my claims that reasons are usually states of affairs and not mental states. I discuss such situations in more detail in Chapter 9, and I argue that if they affect my account of justifiable reasons, they also affect many accounts of objective normative reasons. For now, assume that my intent, and the intent of objective normative reasons theorists, is to explain the reasons that apply to people whose perceptual systems or reasoning skills are not malfunctioning due to drugs, or mental illnesses that induce delusions.[54]

When justifiable reasons and objective normative reasons provide differing accounts of agents' normative reasons, this will normally be because the expectations about the outcomes of an action that it is possible and appropriate for the agent to form when considering what he has reason to do, turn out to be false. The contents of *E*, *V*, and *N* involve estimation and prediction. Even determining what φ is, that is, working out what an action really amounts to, can involve estimation and prediction. The reasoning on which justifiable reason relations are based includes induction as well as deduction. Sometimes the result of even the best inductive reasoning will conflict with reality. In the example above, your circumstances led you to expect that the effect of your taking the medicine will be improved health. Even if you reasoned appropriately given your circumstances, your conclusions about the likely outcome of your action would be false. Recall my claim in Chapter 3 that what makes a state of affairs a reason-explanans is the role it plays within a reason relation. Your doctor's prescribing you a certain medicine is a reason-explanans because of the relationship between this state of affairs and the other components of the reason relation. The state of affairs that would serve as a justifiable reason-explanans for you to act is not a state of affairs that would have served as an objective normative reason-explanans for you to act. That is, if you were able to take a gods'-eye perspective on your situation, you would know that in your particular set of circumstances, your doctor's prescribing you the medicine would not be a good reason for you to take it. If you had an ideal understanding of your situation, the reason relation that describes your set of circumstances would have included a different reason-explanans, perhaps the fact that your doctor prescribed the wrong medicine. This leads to a different, but closely related objection. A critic could concede that reason-explanans are states of affairs, but argue that the reason relations that pick out those states of affairs as reasons are insufficiently connected to reality.

5.2.2 Normative reason relations must be based in reality

Assume that a supporter of objective normative reasons grants that on my account of justifiable reason relations, justifiable reasons are states of affairs, or perhaps quibbles with

---

[54] I stated in Chapter 4 that my account of reasons is intended to apply to people with a wide range of intellectual and practical abilities. I take it that people with what might be considered 'lower' or 'reduced' intellectual and practical abilities have limitations to their functioning, and these limitations will affect which states of affairs are justifiable reason-explanans for them. I don't hold that such agents are malfunctioning in the sense at issue for agents who experience delusions.

me and accepts that they are propositions. Someone who holds that only objective normative reasons are appropriately connected to the actual state of the world may nevertheless argue that merely having a reason-explanans that is a state of affairs is not enough. Such a person might claim that to be normative reasons, reason relations must be appropriately connected to the actual state of the world, and justifiable reason relations are not appropriately connected. Yes, when you have justifiable reason to take the medicine recommended to you by your doctor, your justifiable reason-explanans is that your doctor told you to take it, which is a state of affairs. Nevertheless, the process that picks out that state of affairs as your reason-explanans and determines that you have justifiable reason to take the medicine is flawed, and, perhaps most importantly, subjective, because it depends on your psychological states and capacities.

Objective normative reason relations are sometimes said to be normative because they are based on relevant truths, appropriately ideal deliberation, or both of these ideals.[55] Derek Parfit, for example, argues that what agents have normative reason to do must be determined by considering facts about agents' circumstances (1997, p. 99). Williams argues that correcting an agent's beliefs and reasoning makes normative reasons appropriately normative (1995a, p. 36).[56] And, Smith uses Williams' work as a foundation for his own position when he argues that 'normative reasons are best thought of as truths: that is, propositions of the general form 'A's φ-ing is desirable or required' (1994, p. 95). Justifiable reason relations do not meet any of these criteria for normativity. However, I argue that such arguments do not show that justifiable reasons are not normative.

Williams argues that correcting false beliefs, learning relevant true beliefs, and engaging in appropriate deliberation is necessary for learning what agents have normative reason to do. One reason Williams gives in support of this position is that there should be agreement between third-person accounts of what agents have reason to do and what it is correct for the agent to claim he has reason to do (1981a, p. 103; 1995a, p. 36).[57] If first- and third-person accounts of what an agent has reason to do differ, an agent might be said to have reason to act in two different ways, one derived from the first-person account, one from the third-person account. This will lead to more than just disagreements between agents and observers; often an agent's own account of what there was reason for her to do at a particular time will alter as her perspective changes.

Williams is correct that it is possible to give different accounts of what an agent has reason to do based on the different perspectives from which an agent's situation can be considered. Justifiable reason relations are arrived at by idealisation that is limited to that which is possible and appropriate for a particular individual in a particular set of circumstances. So, in

---

[55] I discuss these claims about the source of normativity in more detail in Chapters 7 and 8.

[56] An alternative reading of Williams' position is discussed in §7.3.

[57] Bernard Williams gives four different arguments to support his position. Two of these are also given by Michael Smith, and one resembles the argument by McDowell given below. Williams' and Smith's arguments are discussed in more depth in Chapter 7.

one sense of the phrase, they might be thought of as reason relations from the first-person perspective. What an agent has reason to do can also be evaluated from a third-person perspective, that is, the perspective of an outsider or of the agent's future self. An outsider's most considered judgement about what an agent has reason to do may differ from what the agent has justifiable reason to do. An outsider may have additional information or reasoning skills that the agent lacks. This is the case even when the outsider is the agent's future self. Note, however, that an outsider's perspective is still limited. The best an outsider can do is work out what he, as an outsider, has justifiable reason to take the agent to have reason to do. Agents' situations can also be evaluated from a gods'-eye perspective, that is, a perspective that involves highly idealised reasoning based on full knowledge of all relevant facts about the world. Although, I argue in Chapter 8 that while gods'-eye perspectives on reason relations are interesting and useful, they are perspectives that obscure, rather than reveal, what real human agents have reason to do.

Assume, for now, that Williams' internalist account of reasons is correct. If so, then from a third-person perspective, it makes sense to say that someone has reason to do something if she would have reason to do it were she aware of the facts or to say that she has no reason to do something if she would not want to do it if she knew certain facts. Consider the situation of an agent looking back on a decision that she made in the past. At the time, after reasoning to the best of her ability, and carrying out any actions she learned were necessary to increase her understanding, she decided that she had overall reason to carry out some action. In the case of the man collapsing next to her, so-called 'appropriate reasoning' may have required only that she quickly move to help. She carries out the action, and as a result learns something new. In this case, she moves to help him, and when the man laughs at her response, she learns that he was faking. Now, once again reasoning to the best of her ability, she decides that she had no reason to act in that way. She concludes, for example, that she had no reason to help the faker.

The description of the agent tricked by someone who fakes a collapse fits the scenarios given by Williams, but it is far from clear that the agent would be correct to conclude that what she had normative reason to do was limited to what she would have taken herself to have reason to do if she had perfect knowledge of all relevant facts rather than the just the evidence available to her. An agent who learns new information about her situation that affects what it is appropriate for her to conclude she has reason to do should alter her judgement about what she has reason to do. Similarly, what an agent ought to be taken to have reason to do from a gods'-eye perspective should take account of the facts. This does not show that justifiable reasons are not normative. In general, if someone collapses next to you, you have good reason to check on that person, and you have good reason to do so even when there is no objective normative reason for you to do so. This means that I am claiming that there is more than one kind of normative reason for action. For many people, this raises the spectre of an agent with competing reasons for action. In §5.3.2, I argue that this is a misunderstanding of what it means to have concepts of both justifiable and objective normative reasons for action.

Michael Smith develops a theory of normative reasons based partly on his acceptance of Williams' claim that reason relations are derived from facts, and partly on his rejection of Williams' semi-Humean approach. According to Smith, for an agent to identify her normative reasons for action, she 'must have no [relevant] false beliefs ... all relevant true beliefs ... [and] must deliberate correctly' (1994, p. 156; 1995a, p. 112).[58] Smith's reference to 'true beliefs' here does not amount to a claim that reasons are true beliefs in a sense that conflicts with my claim that reasons are states of affairs. Smith is giving the conditions under which he holds that a reason relation is true of an agent, not giving an account of reason-explanans. He argues that the information used to determine what an agent has reason to do must not stem from any false beliefs, 'because [the agent] would not have the desire [to act on false beliefs] if he were fully rational' (Smith, 1994, p. 156). Similarly, the information used to determine what an agent has reason to do must be based on 'all relevant true beliefs' because an agent would take herself to have reason to act on all relevant true beliefs 'if she were fully rational' (Smith, 1994, p. 175). This idea of 'full rationality' plays a similar role in Smith's account of 'correct deliberation'. According to Smith, correct deliberation requires that an agent's motivational set is systematically justifiable; that is, the agent's 'underived desires [must] form a maximally coherent and unified desire set' (1994, pp. 156-161; 1995a, pp. 114-117).[59] It is the desires that we would keep if we were fully rational that are the basis for normative reasons. Smith argues that if agents have systematically justified, hence, maximally coherent and unified desire sets, agents' beliefs about what there is reason for any agent to do in a particular situation will converge, and every agent would agree on what an agent has reason to do in some set of circumstances (1994, pp. 164-174; 1995a, pp. 117-125; 1995b, pp. 294-296). Thus, on Smith's approach to practical reason, there are reasons that apply to all agents in virtue of their agency.

Smith's account of normative reasons appears to conflict with my account of justifiable reasons.[60] Like Williams, Smith claims that what agents have reason to do is based on facts that agents will often be unable to access, or that it would make no sense for agents to take themselves to have reason to access. Smith uses the example of someone who wants to buy a Picasso, but doesn't know that there is a Picasso in a local second-hand shop (1994, p. 157). Presumably, it would be inappropriate for him to get an expert in to examine all the paintings in the local second-hand shops to learn whether any of them are Picassos. So, it would not be possible or appropriate for him to take himself to have reason to buy the Picasso in the second-hand shop, and he has no justifiable reason to do so.

---

[58] The word 'relevant' is my insertion, to increase the plausibility of Smith's position, that is, I assume that Smith does not hold that to determine what an agent has reason to do we must assume that he has no false beliefs of any kind.

[59] Smith aims to develop an anti-Humean account of normative reasons, so his account of correct deliberation focuses on the effect that deliberation will have on an agent's desires. Using the terms introduced in §3.2, Michael Smith's prime concern is the effect that deliberation has on what is taken to be the content of *VAL* and *SIG*, rather than the effect that deliberation has on what is taken to be *STATE* or *END*. Although, as explained in earlier chapters, my interest is in the nature of *STATE, END*, and reason relations as such, Smith's account of normative reasons remains relevant here because he gives an account of reason relations based on assumptions about the nature of *STATE* and *END*, not only about the nature of *VAL* and *SIG*.

[60] I argue in Chapter 8 that Smith and I develop accounts of reasons that serve different functions.

Smith examines two different models of a 'fully rational agent': the example model and the advice model (1995a, pp. 110-112). To understand the distinction, imagine two angels sitting on an agent's shoulders: Exemplar and Adviser. Each angel knows everything it needs to know to decide on the best way for an agent to act, and each angel is fully rational. Exemplar tells agents that they have reason to do whatever Exemplar would do in that situation. Adviser tells each agent to do whatever it would be best for someone with that agent's limited rationality to do in that situation. In Smith's example, a short-tempered squash player has just been humiliatingly defeated on the squash court. Should he shake his opponent's hand as he walks off the court? Exemplar says, 'Shake hands'; this is what a perfectly rational agent with a coherent set of desires would do. Adviser knows *A*'s temper is such that if he gets too close to his opponent he will hit him with the racket. So, Adviser says, 'Smile politely at your opponent and leave fast'. Smith writes that '[this] is not something I would be motivated to do if I were fully rational because it is not something I would have any *need* to be motivated to do if I were fully rational' (1995a, p. 111). Smith's Exemplar and Adviser are both tools for determining what an agent has normative reason to do.

Smith acknowledges that Adviser is better at establishing what agents have reason to do because Smith accepts that real agents are not fully rational agents. Although Adviser is a model of a fully rational agent, Adviser accepts that the agent he is giving advice to is not fully rational. (A fully rational agent would presumably have perfect self-control, or would never get cross when beaten on the squash court.) An alternative model of a fully rational agent for justifiable reasons might also be developed. 'Justifier' would be fully rational, but, like Adviser, accept that the agents that she gave advice to were not fully rational. Justifier would recommend that agents took themselves to have reason to act in ways that agents would be justified in taking themselves to have reason to act. Just as Adviser takes account of agents' psychological limitations, Justifier takes account of agents' practical and epistemic limitations. Exemplar ignores any non-ideal features of the agent's character, such as his temper, treating the agent's character as a malleable feature of the situation. Adviser recognises that psychological limitations, such as the beaten squash player's temper, must be taken into account when determining what an agent has reason to do. Justifier limits what an agent has reason to do in a way that reflects agents' practical and epistemic limitations, such as their inability to detect when a painting in a second-hand shop is a Picasso.

The differences between my approach and that of Smith are more profound than the last paragraph may suggest. Smith takes rationality to be foundational and explains reason relations in terms of rationality. Thus, according to Smith, what an agent has reason to do is whatever a fully rational agent would do in that set of circumstances. In contrast, I argue that reasons are foundational and explain rationality in terms of justifiable reasons.[61]

The reasoning that contributes to justifiable reason relations will be limited to what it is possible and appropriate to expect of agents. This means that although agents' reasoning will

---

[61] See Chapter 9.

be justified, it will not ensure that agents have maximally coherent and unified desire sets and will not lead every agent to agree about what an agent in some set of circumstances has reason to do. This does not mean that justifiable reason relations are not normative. The requirement that agents reason in a way that is possible and appropriate for them sets a standard for justifiable reasons based on an achievable level of idealisation. But, I take it that this is, in part, what Smith accepts when he develops his Adviser model of a 'fully rational agent'.

This conclusion seems to result in there being at least two forms of normative reasons, which means that there can be true statements of the form 'he has good reason to do something that he has no good reason to do'. This leaves me vulnerable to a further objection, namely that there cannot be two conflicting forms of normative reason. I consider this objection next.

5.3 Dancy shows justifiable reasons are redundant or incoherent

Two aspects of Jonathan Dancy's work could form the basis for arguments against my account of justifiable reasons. First, in spite of the differences between Dancy's position and mine, Dancy's account of normative reasons may seem very similar to my account of justifiable reasons.[62] Dancy argues that only states of affairs that pass through what he calls an 'epistemic filter' can be normative reasons. I argue that although the states of affairs that pass through Dancy's epistemic filter may be justifiable reasons, Dancy's account of objective normative reasons is significantly different from my account of justifiable reasons. Second, Dancy argues that there cannot be two competing forms of normative reason. Although he is ostensibly arguing that there cannot be both objective and subjective reasons, Dancy's description of such reasons seems to include reasons that agents would be justified in taking themselves to have reason to act on. Thus, Dancy's arguments against the existence of subjective reasons could also be thought to apply to justifiable reasons. However, Dancy's arguments do not show that justifiable reasons are not normative reasons. I show that Dancy is wrong about the consequences of there being two forms of normative reasons.

5.3.1 Dancy's objective reasons are justifiable reasons

In *Practical Reality*, Dancy develops the notion of an 'epistemic filter', which acts as a kind of sieve for reasons (2000, pp. 56-59, 65-66; 2004, pp. 158-159). The epistemic filter separates states of affairs that agents can learn about from states of affairs that agents cannot learn about. Only states of affairs that agents can learn about can become reasons for actions. Dancy gives a good example of this distinction: 'Suppose that, unknown and unknowable to me, someone has been buried alive in my garden during the night. Could this make it wrong of me to go away for a fortnight's holiday?' (2000, p. 57). Dancy argues that it could not, because in this example he cannot know that there is someone buried in the garden. Although Dancy is writing about moral duties here, moral duties arise from moral reasons (2000, pp.

49-50, 59).[63] So, it seems reasonable to assume that Dancy holds that he has no duty not to go on holiday, because he cannot have normative reason to not go on holiday, because he cannot know about the man buried alive in his yard. He writes:

> One can ask what it would have been right to do relative to a certain body of facts, and that body may vary. We may ask our question relative to all the facts, to all the facts available to anyone, to all the facts available to the agent, or any other arbitrarily limited set of facts. (Dancy, 2000, p. 58)

Dancy develops this point, claiming that accounts of what agents have reason to do that are based on sets of facts that are limited in various ways are nonetheless objective (2000, pp. 58-59).

The reasons that agents are said to have for acting once states of affairs are passed through Dancy's epistemic filter are accessible, and in this way they correspond to agents' justifiable reasons for acting. What an agent has justifiable reason to do is relative to facts about what the agent can learn she has reason to do. Dancy's example of a holidaymaker who has a man buried alive in his garden resembles my example of the blind, asnomic man. An agent who felt the need to dig holes in his yard to check for people who had been buried alive every time he went on holiday would be thought to be reasoning in ways that are inappropriate given his circumstances. The states of affairs that he does know about would not justify this sort of investigation into the states of affairs that he does not know about.

Dancy's objective reasons are not, however, the same as my justifiable reasons. On Dancy's account of reasons, justifiable reasons that are not also objective reasons are not normative reasons. Dancy's examples of so-called 'subjective reasons' often seem to correspond to justifiable reasons. Recall the way Dancy runs together 'the world as we take it [to be]' and '[the world as we] defensibly take it … to be' (2009, p. 97). Surely the world as we defensibly take it to be equates to the best understanding we can have of the world? And what agents have justifiable reason to do is partly determined by the best understanding they can be expected to have of the world. One possibility is that Dancy is here referring to, and rejecting, the idea that beliefs about reasons can serve as normative reasons. So, he is rejecting the claim that our reasons are given us by our defensible beliefs about the way the world is, rather than the claim that reasons are given us by those aspects of the world that it would be defensible for us to take to be reasons. However, in the same book in which he introduces the idea of an epistemic filter, Dancy writes:

> [Agents] may do things that are actually ill-advised, but be able to explain why they did these things in such a way as to defend their choice of action, without this showing that what they did was in fact the sensible thing to do. Suppose, for instance, I make a not very sensible choice about what arrangements to make about my pension. And suppose that I can later explain the choice I made by pointing out that there were some crucial facts that I happened quite reasonably to have got wrong, and in this way…exculpate myself. There is a sense of 'justify' in which I can be said to have justified what I did. But this does not

---

[63] The relationship between objective normative reasons, justifiable reasons, and moral reasons is discussed in Chapter 8.

show that the balance of reasons was in favour of the action. It wasn't. Indeed,
the features in the light of which I made my choice turned out all to be a mistake,
and so cannot count even as defeated [pro tanto] reasons. (Dancy, 2000, pp. 6-7)

In this passage Dancy claims that getting some facts wrong due to justifiable error means that 'the features in the light of which … [the choice was made] cannot count … [as] reasons'. This distinguishes Dancy's objective reasons from my justifiable reasons. Justifiable reasons that are not also objective reasons do not get into Dancy's epistemic filter. Only states of affairs that would be taken to be reasons were agents fully informed count as normative reasons.

The relationship between Dancy's epistemic filter, his account of objective reasons and other concepts of reasons is illustrated in the following diagram.

What passes through Dancy's epistemic filter?

Dancy's normative reasons
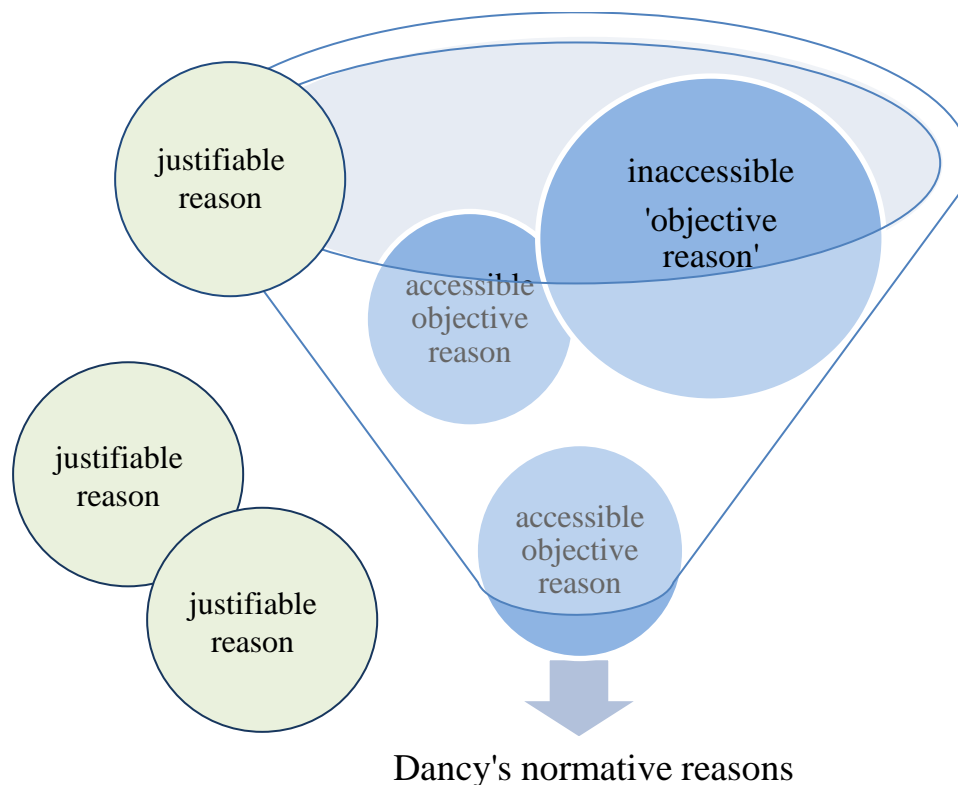
In the diagram above, I label Dancy's objective reasons 'accessible objective reason'. On my account of reason relations, such reasons are states of affairs that are appropriately related to *END, VAL* and *SIG*; that is, they are states of affairs which indicate that acting in a certain way will realise a certain value without conflicting significantly with other values. It is conceivable that all accessible objective reasons are also justifiable reasons because it would be appropriate for agents to learn that they are reasons for action. The sphere labelled 'inaccessible "objective reason"' represents states of affairs that would be objective normative reasons for action if they were accessible to agents. Justifiable reasons that are not also objective reasons are simply not considered normative reasons. This leaves my account

of justifiable reasons vulnerable to Dancy's argument that there cannot be two forms of normative reason.

5.3.2 There cannot be two forms of normative reason

Dancy argues that there is only one kind of reason, the objective kind. Dancy holds that objective, and so normative, reasons are features of the world, while so-called 'subjective reasons' are derived from people's understanding of the world (2000, p. 49). Dancy begins by discussing a situation that resembles a justifiable reasons scenario. You think that the person next to you is in trouble, and this seems to give you reason to help her. However, she isn't in trouble, so there is a sense in which you have no reason to help her (Dancy, 2009, p. 95). He acknowledges that situations such as this lead people to think that there may be both objective and subjective reasons. As he describes them,

> objectivism about reasons [is the claim that] reasons are given us by, or grounded in, features of the situations that we face, not by our beliefs about that situation— not even if we restrict ourselves to our reasonable or permissible beliefs, those we are not at fault in forming … [whereas subjectivism] … holds that our reasons are all given us, not by the world as it is, but by the world as we take it (or defensibly take it) to be. (Dancy, 2009, p. 97, italics in original)

Like most of the other philosophers discussed here, Dancy seems to assume that the difference between subjectivism and objectivism is based on the difference in the ways that the reasons that we take to favour our actions are 'given us'.

When Dancy writes about the way in which 'reasons are given us', he refers to the way in which we end up having reason to act, that is the way in which we end up having reason relations apply to us. When we are given reason to act in a certain way, it means that we have reason to act in a certain way. In other words, it means that we are part of a reason relation. The things that supposedly 'give us' reason to act are what I have been calling reason-explanans. So, when Dancy writes that objective reasons are 'given us by … features of the situations we face', he claims that objective reason relations arise from reason-explanans that are 'features of the situations we face'. Similarly, Dancy argues that reason relations are subjective when they stem from our understanding of the world rather than the world itself, even when that understanding is defensible.

Almost as an aside, I don't like the phrase 'given us'. It implies that the relationship between reasons and agents is passive. Sometimes it seems like this, for example, when you are a so-called 'normal' person and someone next to you suddenly clutches his chest and collapses at your feet. At other times reasons are anything but given. The features of a situation may only become visible as reasons after significant mental or physical struggle, or an extended period of training of some kind. In spite of his words, it seems likely that Dancy accepts that reasons are sometimes uncovered after great effort rather than 'given'. However, this seemingly minor point does direct attention to one of the key differences between certain forms of objective reasons and my concept of justifiable reasons. The method through which agents reach conclusions about what they have reason to do sometimes seems so unimportant in

accounts of objective normative reasons that it is barely mentioned, or not mentioned at all. In contrast, when justification is an essential part of what it means for an agent to have reason to act, the appropriate process for deciding what there is reason to do is what matters most.

Mark Schroeder offers a different interpretation of Dancy's position, based on Dancy's earlier work and personal communications. According to Schroeder, Dancy accepts that there are both subjective and objective reasons. Schroeder mentions this when he offers Dancy as an example of someone who holds that subjective and objective reasons are reasons in virtue of their relationship to some third thing (Schroeder, 2008c, p. 15, fn 23). Writing about the relationship between subjective and objective reasons, Schroeder, writes:

> [Dancy] thinks that neither objective reasons nor subjective reasons are basic, but that both objective and subjective reasons are accounted for in terms of some third thing: 'there are just two questions which we use the single notion of a reason to answer.' (Dancy, 2000, p. 2, cited in; Schroeder, 2008c, p. 15, fn 23)

Dancy uses 'reason' to refer to 'reason-explanans, that is, considerations that favour, or are taken to favour, acting some way. In the quoted sentence, the 'two questions' to which Dancy refers are questions about, 'his reasons for doing it' and whether there was 'any reason for doing it at all' (2000, p. 2, italics removed). In other words, Dancy is distinguishing between motivating and normative reasons, not subjective and objective reasons.

Dancy holds that motivating and normative reasons are reasons in virtue of their relationship to a third reason concept because 'one and the same reason can be both motivating and normative', a thesis he argues for over the course of *Practical Reality* (2000, p. 6). According to Dancy,

> reasons are given us by features of the situation, rather than by our own psychological states—unless those states function merely as features of the situation. Our reasons … are objective rather than subjective or relative to our psychology. (Dancy, 2000, p. 157)

When Dancy writes here that motivating and normative reasons are both objective, he means, as I wrote in §3.3, that both motivating and normative reason-explanans are states of affairs. This passage agrees with the argument of Dancy's from 'Rationality and Reasons' that is described above; the same things serve as motivating or normative reasons, and there are no subjective reasons (2009, p. 97).

Dancy's argument that there are only objective reasons is as follows (2009, pp. 97-99). There are three possible positions on objective and subjective reasons: there could be only objective reasons, only subjective reasons or both objective and subjective reasons. Dancy takes it as obvious that the way that the world is gives agents reasons to act in particular ways, and argues that, if this is so, there are definitely objective reasons.[64] So, he rejects any position that claims that there are only subjective reasons, and concludes that if there are subjective reasons, then there must be both objective and subjective reasons. Next, Dancy considers the

---

[64] In 'Rationality and Reasons' Dancy takes this as obvious, however his book *Practical Reality* is an extended argument for this thesis (2000, 2009).

possibility of a dualist position according to which agents should 'first, do what [they] have most objective reason to do, and, second, do what [they] have most subjective reason to do' (2009, p. 97). Dancy argues that this form of dualism about subjective and objective reasons is peculiar. So-called 'subjective' and 'objective' reasons can conflict. If the person next to you fakes a collapse, you have a subjective reason to help him and an objective reason not to help him. However, individual agents can't tell whether the reasons whose pull they feel are subjective or objective. According to Dancy, only outside observers can identify and distinguish between an agent's objective and subjective reasons. So, individual agents cannot be in a position to weigh subjective and objective reasons against each other and choose between them. Thus, Dancy concludes, there cannot be both sorts of reasons, so there can only be objective reasons.[65]

Dancy is correct that the form of dualism he describes is peculiar. As he writes, it is not possible for agents to distinguish between subjective reasons and objective reasons. Nor would it be possible for agents to distinguish between justifiable reasons and objective reasons. So, it would be absurd to advise agents to consider their objective reasons, consider their subjective or justifiable reasons, and then decide what to do. The situation is even worse than Dancy describes. If what an agent has objective normative reason to do is derived from a highly idealised perspective, real human observers will also be unable to determine whether subjective reasons, or my justifiable reasons, are also objective reasons. A third party's identification of an agent's objective, subjective, or justifiable reasons is nothing more than that observer's best interpretation of the agent's situation. Even observers cannot distinguish between their subjective account of what an agent has reason to do and what the agent has objective reason to do.

The theory of justifiable reasons that I develop differs significantly from the form of dualism about reasons that Dancy rejects. Dancy argues that there must be objective reasons because we know that the way the world is gives us reason to act (2009, p. 97). I agree with Dancy that the way the world is gives us reason to act, hence my claim that reason-explanans are states of affairs. On my account of reasons, Dancy's claim that the way the world is gives agents reasons to act leaves open three possibilities: there are only objective reasons; only justifiable reasons; or both objective and justifiable reasons. Given that highly idealised accounts of objective reasons cannot explain what real agents have reason to do, or explain our conceptions of self-regulation, rationality, praise- and blame-worthiness, I would not conclude that there were only objective normative reasons without seeing a strong argument that this was the case. I believe I can give an account of justifiable reasons that avoids objections directed against positions such as mine. However, I do not claim that justifiable reasons are the only form of normative reason. I support a form of dualism about reasons, but not one that succumbs to Dancy's objection.

Justifiable and objective normative reasons serve different purposes. There are no occasions where agents compare, or need to compare, their justifiable and objective normative reasons

---

[65] Torbjörn Tännsjö makes a similar claim for similar reasons (2010, pp. ix-x).

for acting. Justifiable reasons are the reasons that real human agents ought to be disposed to take themselves to have reason to act on and ought to be disposed to use to assess their own actions. Justifiable reasons are also the only normative reasons that real human agents can use to assess other people's actions. In contrast, objective normative reasons serve an important role within philosophical theorizing, both inside and outside the discipline of philosophy. [66] Sometimes justifiable reasons are also objective reasons, but objective reasons only become practical reasons for real human agents when they are also justifiable reasons. These are not competing reasons for action; for two things to compete, they need to be in the same contest.

5.4 Conclusion

This chapter considered four possible objections to justifiable reasons, and concluded that they fail to show that justifiable reasons cannot provide a coherent theory of normative reasons for action.

Railton argues that first-personal accounts of reasons cannot successfully describe agents' reasons for acting because agents who act appropriately in response to reasons often have a limited understanding of the normative reasons that guide their actions. I showed that even though Railton's description of first-personal accounts of reasons resembles my account of justifiable reasons, my account of justifiable reasons breaches first and third personal accounts of reasons in the way that Railton argues is required of successful accounts of normative reasons.

A number of people claim that if they are to be normative, reasons must be objective. The agent-relativity of justifiable reasons may seem to make such reasons subjective, thereby calling their normativity into question. However, arguments that reason-explanans must be real or that reason relations must be based on reality do not show that justifiable reasons are subjective in a way that prevents them from being normative reasons. I take it that the reasons that serve as the basis for normative reason relations are of a kind, and I agree that they are usually states of affairs, not psychological states. If claims about subjective reasons are claims that there is a specific variety of grounds for action, a 'subjective reason', and that subjective reasons are psychological states, then I agree with Wallace and Dancy when they claim that there are no subjective reasons. However, justifiable reasons do not fit this characterisation of subjective reasons. The reasoning that picks out justifiable reason-explanans as reasons, and from which justifiable reason relations are derived, is limited by agents' abilities. However, this does not mean that justifiable reason relations are subjective in a way that undermines their normativity. I also showed that justifiable reasons and objective reasons do not compete in a way that makes the existence of two forms of normative reasons problematic. Objective normative reasons may be the reasons that agents would have reason to act on if they were capable of a gods'-eye perspective on the world, but justifiable reasons are the normative reasons that agents act on.

---

[66] The function of objective normative reasons is discussed in Chapter 8.

John McDowell nicely expresses what is required for a theory of reasons to be normative:

> Reason-giving explanations require a conception of how things ideally would be sufficiently independent of how any actual individual's psychological economy operates to serve as a basis for critical assessment of it. In particular, there must be a potential gap between the ideal and the specific directions in which a given agent's motivations push him. (McDowell, 1995, 76)

Although McDowell's comments are directed at claims about the relationship between normative reasons and agents' motivations, his criteria are just as relevant for assessing the relationship between normative reasons and agents' practical and epistemic circumstances. Just as an account of reasons must allow for a potential gap between ideal and actual motivations for an agent's reasons to be normative, there must be a potential gap between the ideal and actual nature of agents' understanding of their practical and epistemic circumstances for the agent's reasons to be normative. The role that justification plays in determining what an agent has justifiable reason to do takes account of agents' psychological states in a way that leaves room for critical assessment, and in doing so it leaves a gap between what an agent has justifiable reason to do and the agent's motivating reasons that ensures that justifiable reasons are normative. One way to get this gap between the ideal and the actual is to hold that an agent only has reason to act in a certain way if he knows all the relevant facts and has reasoned about them perfectly. This gap between real agents and objective normative reasons will often be insurmountable. But, the gap between agents and what they would be justified in taking themselves to have reason to do is both normative and surmountable.

These four objections are not the only potential objections to my account of justifiable reasons. Philosophers who argue in support of so-called 'internal reasons' make claims about reasons and normativity that appear to conflict with some of my claims about justifiable reasons, and I consider these at the end of Chapter 7. My account of reasons could also be considered vulnerable to objections to subjective reasons. So, in the next chapter I show that my account of justifiable reasons differs from accounts of subjective normative reasons that seem to have been developed to explain situations that resemble the examples that I use to support the development of justifiable reasons.

6. Subjective Normative Reasons

6.0 Introduction

Current work on reasons for action focuses primarily on 'objective normative reasons', with subjective normative reasons occasionally mentioned as an alternative. There is no agreed upon single concept of objective reasons, but the term is usually used to refer to reasons for action that lead to a correct outcome, realise certain values, or stem from ideal reasoning based on all the relevant facts.[67] Reasons of this kind link reasons and values in a way that facilitates certain outcomes within value theory, meta-ethics and philosophy of action, and examples of this are given in Chapter 8. Unfortunately, objective normative reasons often fail to correspond to real people's reasons for acting. Thus, work on objective normative reasons is often irrelevant to issues such as appropriate self-regulation, praise, blame and rationality. Some philosophers introduce the concept of subjective normative reasons in response to this problem. However, subjective normative reasons are usually defined in terms of agents' beliefs, and, as agents' beliefs can be false, this makes it difficult to see how subjective normative reasons can be good reasons for acting. I argue that the drawbacks associated with the notion of subjective normative reasons are a barrier to this concept successfully explaining the intuition that people can have normative reasons to act in ways that run counter to the objective reasons they have for acting. This forms part of my justification for introducing a new concept of reasons, 'justifiable reasons', which captures the useful element of subjective normative reasons, but avoids the drawbacks associated with subjective reasons.

Return to the bus example: You are on a bus, worrying that your stuffy nose means that you have a head cold. Just before your stop, the person next to you collapses to the floor. Although you cannot know this, he is faking. If you help him, he will continue to act in this way, which will cause many more upsets than benefits. What do you have reason to do?

On the accounts of objective normative reasons discussed in this thesis, you have no reason to help the collapsed person. If you knew all the relevant information and reasoned appropriately, you would realise that he did not need help. If consequences determine what you have objective reason to do, better consequences follow from not helping him than from helping him. There is, however, no way that you can know that he is faking and no way that you can know the consequences of helping him. So, there is a sense in which objective normative reasons have no relevance to how you should act.

Whatever you would have reason to do if you had some form of ideal understanding, if you are a normal human agent, you are likely to think that if someone collapses next to you, you

---

[67] Some claim that Williams' internalism can be seen as an account of objective normative reasons (Joyce, 2001, p. 53; Railton, 2008, p. 238; Schroeder, 2008c, p. 13, fn 18; Setiya, 2004, p. 271). However, David Sobel argues that Williams' internalism is a subjectivist theory (2001, pp. 473-474). This issue is irrelevant here.

have reason to help. Some philosophers argue that when the correct response of a real agent differs from the correct response of an ideal agent, the agent has conflicting reasons for action. On this view, when the person next to you collapses in a heap on the floor, you have no objective normative reason to help him, because he doesn't really need help. Such reasons are called 'objective' because they are based on reality. However, you do have a subjective normative reason to help him, because if the information available to you tracked the truth, this would be the correct response to that information. These are called 'subjective reasons' because they are based on agents' beliefs about what they have objective reason to do.

Current analyses of subjective normative reasons tend to be weak, undeveloped introductions. When subjective normative reasons are defined as 'beliefs about objective reasons', this leads people to think that objective reasons are primary and focus predominantly on objective reasons. Hence, Cullity and Gaut write, 'in seeking an account of normative practical reasons, it is objective normative reasons that will be our primary concern: from this an account of subjective ones will follow' (1997, p. 2). Subjective reasons are not mentioned again. Similarly, Mark Schroeder writes that objective normative reasons are fundamental, and rarely mentions subjective reasons after the first chapter of his book (2008c). The analysis of subjective reasons as 'beliefs about…' emphasises the supposed subjectivity of the concept, which makes it difficult to see how the account refers to normative reasons. It seems to have this effect on people's understanding of such reasons even when the theory of subjective normative reasons that is developed is not a theory about agents' beliefs, but rather a theory about what agents would believe if they were to engage in appropriate reasoning. More importantly, the examples used in arguments about subjective reasons suggest that they are intended to capture real agents' reasons for acting, but the reasons that real agents have for acting are not beliefs about objective reasons, or so I will argue.

In this chapter I show that although accounts of subjective normative reasons appear to be motivated by the inability of objective normative reasons to connect with real agents' decision making, these accounts fail to achieve that goal. The accounts of subjective normative reasons described in §6.1 are not 'normative reasons' in the usual sense of the term. They seem instead to be accounts of reasons where agents have beliefs that it would be irrational for them not to act on given their goals. Maureen Sie, Marc Slors and Bert van den Brink, whose position is described in §6.1.1, seem partly motivated by a desire to explain situations where real agents' reasons do not correspond to ideal agents' reasons (2004). However, unlike my concept of justifiable reasons, the concept of subjective normative reasons developed cannot serve the required purpose. Mark Schroeder also uses examples where an agent's understanding of the world is flawed to explain his concept of subjective normative reasons (2008b, 2008c). His position is outlined in §6.1.2. Schroeder's subjective normative reasons need not motivate agents. Instead Schroeder seems to use the term to refer to an agent's having reason to take the means to his ends. In §6.2.1 I describe the account of subjective normative reasons given by Cullity and Gaut and in §6.2.2 that given by Richard Joyce (Cullity & Gaut, 1997; Joyce, 2001). Although these conceptions of subjective normative reasons resemble my concept of justifiable reasons, on these accounts such reasons amount to beliefs about objective reasons; in §4.4 I argued that this is a mistake, and in §6.3 I

expand on this argument. Finally, in §6.4, I consider Peter Railton's account of what he calls 'subjective reasons' (2008). Railton's subjective reasons are normative, but he presents an internalist account of reasons, which makes his account significantly different from mine.

6.1 Unjustified subjective normative reasons

According to Sie, Slors, and van den Brink, and Schroeder, justification plays no role in determining whether something is a subjective normative reason. So, although the examples they use initially suggest their accounts may resemble my account of justifiable reasons, this similarity is illusory. On the accounts of Sie, Slors, and van den Brink, and Schroeder, an agent who wants to get drunk and believes that beer becomes more alcoholic when you dilute it with water has a subjective normative reason to add water to her beer before drinking it. Presumably this subjective form of reason is called 'normative' because an agent with such goals and beliefs might be thought irrational if she did not act on her false beliefs.

6.1.1 Sie, Slors, and van den Brink

The account of subjective normative reasons given by Sie, Slors, and van den Brink begins with examples that suggest that they are reasons that agents would be justified in taking to be normative reasons for their actions. Their function appears to be to explain situations where the reasons that real agents would be justified in acting on differ from those that would apply to idealised agents. Yet, their account of reasons cannot serve this function.

Sie, Slors and van den Brink introduce their distinction between subjective and objective normative reasons with the following example:

> [I am ill and visit my doctor for help.] My doctor might know … that I have objective reason to take the red pills instead of the blue ones because … the red ones … contain drug *D* that will give me back my long-desired health. But when I [firmly] believe that the blue pills rather than the red ones contain *D*, then, knowing that *D* promotes my health, I have a subjective normative reason to take the blue ones. That is I [firmly] believe that I ought … to take the blue pills. Thus, my subjective reason is my *belief* that I have an *objective* reason to take the blue pills. It is subjective since it is just *my* belief which can be, and in this case is, false. (Sie et al., 2004, p. 4, italics in original, 'firmly' substituted for 'truly')

So, Sie et al. use 'subjective normative reasons' to refer to beliefs about objective reasons, and the agent's belief is the reason that explains what the agent has subjective normative reason to do. [68]

When subjective normative reasons are taken to be beliefs, it is difficult to see how such reasons, or reason relations, can be normative reasons. Sie et al.'s 'subjective normative

---

[68] If an agent who believes that she ought to act in some way is always to some degree motivated to act in that way, then Sie, Slors and van den Brink's account of subjective normative reasons is also an account of motivating reasons.

reasons' are not normative in the sense that normative reasons are usually taken to be normative, but the implication seems to be that the agent would exhibit means-end irrationality if she failed to act on them. For example, if I want to be cured, my firm belief that the blue pills are the ones that will cure me is a reason for taking me to be acting irrationally if I don't take the blue pills. It seems appropriate for the agent to be at least somewhat motivated to act according to such reasons.

Sie et al. qualify their account of subjective normative reasons in a way that makes the strangeness of calling these reasons 'normative reasons' clear:

> The crucial thing here is that even subjective normative reasons are eligible for justification only when they are rightfully thought to be objective reasons *as well*, i.e., only when the subjective *belief* that this subjective reason is an objective reason as well is *true*.' (Sie et al., 2004, p. 4, italics in original)

So, an agent has a subjective normative reason for acting some way when she believes that she has an objective reason to act in that way. However, the agent's subjective normative reason is only a normative reason when the agent's belief that she has reason to act in that way is correct, that is, when the agent has an objective reason to act in that way. On this account, an agent who believes that she has reason to take the blue pills when she has no objective reason to do so, does not have a normative reason to take the pills.

Sie, Slors and van den Brink's example resembles the situation you are in when the person next to you collapses on the bus and you cannot detect that he is faking. Their example appears even more like mine if the example is altered to state that my usually trustworthy and knowledgeable doctor told me to take the blue pills. Arguments from authority can be a justifiable method for deciding what to do, so if my doctor tells me to take the blue pills, I probably have good reason to swallow the blue pills. However, Sie et al.'s concept of subjective normative reasons does not describe a form of reasons for actions that agents would be justified in taking to be reasons.

Sie, Slors, and van den Brink claim that an agent's subjective normative reason for acting in some way is only a normative reason when it is also an objective reason because '[as] soon as I know that my belief about the blue pills…is false, I would not consider my subjective reason for taking the drug justifiable' (2004, p. 4). Sie et al. are correct that if an agent learns that her beliefs about how she has reason to act are wrong, she would not be justified in concluding that she has reason to act as she initially supposed she had reason to act. However, I argue that an agent may learn that she did not have an objective reason to act in some way without concluding that her belief that she had reason to act in that way was unjustifiable.

An agent who learns that her decision to act in a certain way led to an unintended outcome may correctly conclude that her decision was unjustifiable. To see this, imagine a situation where an agent decides she has reason to act some way for a silly reason. Assume that the agent in Sie et al.'s example was told by her doctor to take the red pills, but in spite of this she decided to take the blue pills because she felt whimsical and blue is her favourite colour. Afterwards, clutching whichever part of her the drug was supposed to heal, she feels foolish

and realises that her decision was unjustified. In such a situation, the agent had a motivating, but not normative, reason to take the blue pill. So, as Sie et al. suggest, an agent who learns that her beliefs about what she had reason to do were mistaken may correctly conclude that she did not have a normative reason to act in that way. It does not follow from this that an agent who learns that she had a mistaken belief about what she had reason to do should conclude that her belief was unjustified.

An agent may learn that her decision about what she had reason to do was based on a mistaken interpretation of her situation and still insist that her reason for taking the pill was justifiable. If the agent decided that she had reason to take the blue pills because her well-qualified, trustworthy doctor assured her that this was the case, she has no grounds for believing that she was unjustified when she took the pill, and she had a justifiable reason to do so. This is the case even though she would certainly conclude that she had no reason to take the blue pills after she learned that her doctor made a mistake.

Sie, Slors, and van den Brink do not develop an account of subjective normative reasons that describes the reasons that agents have when their understanding of the situation that they are in justifiably differs from an idealised understanding of their situation. Nor do their arguments about the nature of subjective normative reasons show that it is impossible to develop an account of normative reasons that applies to agents who have a partial and potentially mistaken understanding of the world or the consequences of their actions.

6.1.2 Schroeder

Mark Schroeder introduces his notion of subjective normative reasons as part of a larger argument in support of Humeanism about reasons (2008c). At first glance, Schroeder seems to be suggesting that we need a concept of reasons that allows for human fallibility but retains normativity. He argues that there are three forms of reasons: 'motivating reasons', which he interprets in the standard way; 'subjective reasons', based on 'what the agent believes independently of how things actually are'; and 'objective reasons', based on 'how things are independently of the agent's beliefs' (2008c, pp. 13-14).[69] Schroeder uses Bernard Williams' classic example: Bernie wants a gin and tonic and is holding a glass that he believes contains just that. Unfortunately, it contains gasoline (Schroeder, 2008c, p. 13). Like Sie, Slors, and van den Brink, Schroeder claims that Bernie has a subjective normative reason to act on his belief. In fact, Schroeder claims that we would criticise him if he did not drink from his glass. Bernie is also said to have an objective normative reason not to act on his belief (Schroeder, 2008c, p. 13). Schroeder's subjective normative reasons bridge motivating and normative reasons, but need not motivate. As with Sie et al., Schroeder's concept of subjective normative reasons differs from my concept of justifiable reasons because there is no

---

[69] Schroeder also discusses subjective and objective reasons in 'Having Reasons', but does not call them subjective normative reasons in this article (2008b). The account presented here is based on his book, *Slaves of the Passions* (2008c).

requirement that an agent be justified in order to have a Schroederian subjective normative reason.

Schroeder argues that objective normative reasons are the most basic kind of reason and explains both subjective normative reasons and motivating reasons in terms of them (2008c, pp. 13, 15; 2009, p. 13). Schroeder initially describes objective normative reasons in a way that suggests that they can be understood as instrumental means to currently desired ends.

An agent has a desire.	Bernie wants a gin and tonic.
Some reason is an accurate sign that acting in a particular way is a means to an end that satisfies that desire.	The gin in the bottle is a sign that pouring the contents of the bottle into a glass and adding tonic will make a gin and tonic.
The agent has an objective normative reason to act in that way.	Bernie has reason to pour the contents of the bottle into a glass and add tonic.

This looks a lot like a prototypical account of Humeanism about reasons, but Schroeder's account of objective normative reasons is more sophisticated than this. One of the classic criticisms of Humeanism about reasons is that agents can have desires that they ought not to have, or fail to desire things that they ought to desire. If Bernie is an alcoholic, then Bernie's desire for a gin and tonic does not give him reason to mix the gin with tonic and drink it. Similarly, whatever the alcoholic Bernie desires, he ought to desire not to drink a gin and tonic. Schroeder develops several ways of dealing with the possibility that desires can be misplaced or have mistaken strength (2008c).[70] If Bernie's situation is described in such a way that if he were to think correctly about his situation, he would not desire a gin and tonic, then Bernie does not have reason to take the means to satisfying his uncorrected desire that he wants a gin and tonic. The relevant point is that there is an objective normative reason for an agent to act in some way only when he would have reason to act in that way if his desires met appropriate ideals. Picking out features of the world that are reasons in the sense that they show that if he takes certain means he will achieve ends he currently desires is not enough to make those features of the world reasons for him to act.

Schroeder explains subjective normative reasons in terms of objective normative reasons. Using 'R' to refer to the agent's reason-explanans, 'X' to refer to the agent, and 'A' to refer to the relevant action, he writes:

> For R to be a subjective reason for X to do A is for X to believe R, and for it to be the case that R is the kind of thing, if true, to be an objective reason for X to do A. (Schroeder, 2008c, p. 14)[71]

---

[70] See, in particular, Chapters 5-8 of Schroeder's *Slaves of the Passions*, but this argument is ongoing throughout the book (2008c).

[71] A slightly different account is given in section 2.1 of Schroeder's book, but the revision is associated with the role of desires and is irrelevant here (2008c).

Although Schroeder uses the term 'subjective reasons' rather than 'subjective normative reasons', the context shows that this quote refers to subjective normative reasons. So, according to Schroeder, an agent has a subjective normative reason if the content of a belief of his would be an objective normative reason if the content of that belief were true.

Consider Schroeder's account of the relationship between subjective normative reasons and objective normative reasons using his example, with reason 'R' treated as the proposition that the glass contains gin and tonic. Bernie wants a gin and tonic. Given Bernie's situation, if there is gin and tonic in his glass, he has an objective normative reason to drink from the glass. On Schroeder's analysis, if Bernie believes that the glass contains gin and tonic, then, because he would have an objective reason to drink from the glass if his belief were true, he has a subjective normative reason to drink from the glass.

After explaining subjective normative reasons in terms of their relationship to objective normative reasons, Schroeder explains motivating reasons in terms of subjective normative reasons:

> For R to be the (motivating) reason for which X did A is for the fact that R was a subjective normative reason for X to do A to constitute an explanatory reason why X did A. (Schroeder, 2008c, p. 14)

On this account, something is a motivating reason only when it is a subjective normative reason that explains why an agent acted as she did. Return to Bernie in Schroeder's example. Bernie wants a gin and tonic. Bernie drinks from a glass because he believes it contains gin and tonic. For Bernie's belief that the glass contains gin and tonic to be the motivating reason for his action, it needs to be the case that if the glass contained gin and tonic, Bernie would have an objective reason to drink from it.[72]

This way of using subjective reasons to link motivating reasons to objective reasons makes Schroeder's account of motivating reasons inappropriately restrictive, that is, it entails that certain things that are usually taken to be motivating reasons are not considered to be motivating reasons. Accept for now that the contents of beliefs can be motivating reasons, and that Bernie drinks from the glass because he believes it contains gin and tonic. Given our assumptions, Bernie's belief that the glass contains gin and tonic is his motivating reason. Now assume that Bernie is an alcoholic who desperately wants to stop drinking. If he is an alcoholic, then arguably he has no objective reason to drink alcohol, and so no objective

---

[72] Schroeder rarely mentions subjective normative reasons in the rest of this book. Schroeder's principle goal here is to develop an argument in support of the Humean theory of reasons (2008c). Schroeder claims that if his accounts of subjective and motivating reasons are correct, 'then the fundamental sense of "reason" that is relevant in all but purely explanatory contexts is the objective normative sense' (Schroeder, 2008c, p. 15). For that reason, he focuses on objective normative reasons for the remainder of his book. Broadly speaking, Schroeder argues that the reasons that agents have reason to act on are objective normative reasons, but he is subjectivist about reasons in the sense that he believes that it is agents' psychological states that are responsible for agents having reason to act (2008c, pp. 1-6, 11-12, 192-197). As mentioned, the Humean theory of reasons is not directly relevant to my account of justifiable reasons, so rather than discuss Schroeder's overall project in any detail, my comments concentrate on his account of the relationship between motivating, subjective and objective reasons.

reason to drink gin and tonic. In situations like this Schroeder's account entails that if Bernie drinks from the glass because he believes that the glass contains gin and tonic, his belief that the glass contains gin and tonic is not a motivating reason, because even if Bernie were right that the glass contained gin and tonic he would still have no objective reason to drink from the glass. Yet, Bernie's belief that the glass contains gin and tonic was his reason for drinking from it, so surely this is his motivating reason. This problem with Schroeder's account of motivating reasons is only semi-relevant here. However, this same problem affects Schroeder's account of subjective normative reasons.

Schroeder's account of subjective normative reasons is also too restrictive, and restrictive in a way that makes one aspect of the difference between my account of justifiable reasons and Schroeder's account of subjective normative reasons clear. Schroeder claims that for an agent to have a subjective normative reason to act, the truth of the reason-explanans, *R*, must provide an objective reason for the agent to act. This way of understanding subjective normative reasons allows for the following situation:

X desires to do *P*.	Bernie wants a gin and tonic.
X believes *R*.	Bernie believes that his glass contains gin and tonic.
R is true.	Bernie's glass contains gin and tonic.
R is **not** the kind of thing, if true, to be an objective reason for *X* to do *A*.	The fact that his glass contains gin and tonic is not an objective reason for Bernie to drink from the glass.
R is **not** a subjective reason for *X* to do *A*.	The fact that his glass contains gin and tonic is not a subjective reason for Bernie to drink from the glass.

This situation may at first seem exactly what Schroeder requires. The problem is that in this example *A* is a means to *P*, so Schroeder is committed to saying that even though the agent knows that *R* means that *A* is a means to achieving one of his ends, the agent has no subjective reason to do *A*. This situation is possible because the relationship that Schroeder draws between subjective and objective normative reasons means that the truth of an agent's beliefs about *R* will not give that agent an objective reason to act if the agent was mistaken about the appropriateness of his goal.

I mention this problem with Schroeder's position not just because it is a problem for his account of subjective normative reasons; it also helps distinguish subjective normative reasons from my account of justifiable reasons. Imagine that Bernie is at a point in his life where he has yet to fall into alcoholism, and would lose his desire to drink alcohol if he knew that doing so would bring him years of grief. Nevertheless, in his faultless ignorance, Bernie has a justifiable, but mistaken, desire to drink a gin and tonic. Bernie justifiably and correctly

believes that there is gin and tonic in his glass. But the gin and tonic in his glass is not an objective reason for Bernie to drink from the glass, because if Bernie had idealised knowledge about himself and his circumstances he would have no desire for a gin and tonic. Yet, although Bernie does not have a Schroederian subjective normative reason to drink from the glass, if he has reasoned in a way that is appropriate given his circumstances, he would be justified in concluding that he has a normative reason to drink from his glass.

The relationship Schroeder draws between objective, subjective and motivating reasons is also too promiscuous, and promiscuous in a way that also distinguishes his account of subjective normative reasons from my account of justifiable reasons. It is difficult to see why Schroeder's 'subjective normative reasons' are normative. According to Schroeder, Bernie has a subjective normative reason to drink from his glass, just because, if his belief that it contains gin and tonic were correct, he would have an objective normative reason to drink from the glass. On this conception of reasons, Bernie could have ridiculous reasons for thinking that there is gin and tonic in the glass, and still be said to have a subjective normative reason to drink from the glass. Thus, Schroeder's subjective normative reasons exist in situations where agents would be unjustified in acting on them, unlike justifiable reasons.[73]

It seems odd to call subjective normative reasons 'normative' when there is no requirement that such reasons are justified. Schroeder doesn't call such reasons 'normative' in two articles on the distinction between objective and subjective reasons, and he does not explain why they are normative in *Slaves of the Passions* (2008b, 2008c, 2011). Presumably he takes them to be normative reasons because agents would exhibit means-end irrationality if they failed to act according to such reasons. If this is correct, Schroeder's account of subjective normative reasons resembles that of Sie, Slors and van den Brink.

Schroeder's subjective normative reasons are incapable of explaining the kind of reasons we take real agents to have in situations where their grasp of their situation is limited. His account does not require that subjective normative reasons be justified. Instead, he uses the term to refer to situations where people happen to believe a proposition that, if true, would mean that acting in a certain way would achieve an appropriate goal. Furthermore, on Schroeder's conception of reasons, agents with justified false beliefs about their goals have no subjective normative reason to pursue those goals.

6.2 Justified subjective normative reasons

Not all conceptions of subjective normative reasons fail to be normative. Unlike Sie, Slors and van den Brink, and Schroeder, the accounts of subjective normative reasons given by Garrett Cullity and Berys Gaut, and by Richard Joyce require such reasons to be justified. In this way Cullity and Gaut, and Joyce take a step towards being able to account for the

---

[73] In a related paper on epistemology, Schroeder argues that subjective normative reasons need not be justified (Schroeder, 2011).

seemingly normative reasons for acting that agents have when they have a justifiable but somehow mistaken understanding of the relevant circumstances. However, both Cullity and Gaut, and Joyce, define subjective normative reasons in terms of what agents would be justified in believing they have objective reason to do. This way of describing subjective normative reasons is misleading. The history of analyses of reasons means that philosophers often react to any explication of a reasons concept in terms of 'beliefs' with the thought that the concept is dependent on psychological states. However, neither Cullity and Gaut nor Joyce claim that subjective normative reasons are beliefs. In both cases, the word is used as part of a counterfactual claim of the kind 'if A were to reason appropriately, A would believe that she had an objective normative reason to φ.' More importantly, the claim that subjective normative reasons have the relationship to objective reasons suggested by Cullity and Gaut, and Joyce, does not allow their conceptions of subjective normative reasons to account for the kinds of examples they intend them to account for.

6.2.1 Cullity and Gaut

Cullity and Gaut introduce the concept of subjective and objective normative reasons with an example similar to that of Sie, Slors and van den Brink:

> Suppose your doctor tells you to take a certain medicine, but this happens to be a mistake, and it will harm you. What should you do?.... There is a clear sense in which *that your doctor has told you to do so* is a normative reason to take the medicine. However, there is an equally clear sense in which *that it will harm you* is a normative reason not to take it. We might call this the distinction between subjective and objective normative reasons, respectively. (Cullity & Gaut, 1997, pp. 1-2, italics in original)

This concept of subjective normative reasons, reasons that you have because 'your doctor told you to do so', initially seems to be a concept of reasons that agents would be justified in taking to be normative reasons for their actions. Cullity and Gaut treat the proposition '*that your doctor has told you to do so*' as the reason-explanans for acting, not the agent's belief. This bases the agent's reason for acting on the state of the world rather on the agent's psychological state. Just as importantly, this reason-explanans appears to justify you concluding that you have reason to take the pill. This introduction of justification as an important element of an account of subjective normative reasons distinguishes Cullity and Gaut's account from those of Sie et al., and Schroeder. If all is as it should be, people are justified in accepting what their doctors tell them about medicine. If this were not the case, you would not have a normative reason to act on the doctor's advice. Cullity and Gaut seem to assume that you can have normative reason to follow that advice even when the doctor is wrong. In this way, their account resembles my account of justifiable reasons.

Like the other authors whose positions I consider, Cullity and Gaut claim that there is a 'clear' relationship between objective and subjective normative reasons: 'I have a subjective normative reason to φ whenever I am justified in believing that I have an objective normative reason to φ' (1997, p. 2). This may seem an odd claim; even philosophers are unlikely to think, 'I believe that I have an objective normative reason to act in this way' when they

decide to follow their doctor's instructions. However, presumably an agent would be justified in believing that she has an objective reason to act in some way if she were disposed to agree to four statements. First, that she has an accurate understanding of the relevant facts of her situation. Second, that she has reasoned appropriately. Third, that her goals are appropriate. And, fourth, that she has an accurate understanding of the outcome of the proposed action. So, I assume that claims that an agent has a subjective normative reason to act some way whenever she is justified in believing that she has an objective normative reason to act in that way amount to claims that an agent has a subjective reason to act in some way whenever she would be justified in believing that these four statements are true of her. This relationship that Cullity and Gaut draw between objective and subjective normative reasons means that their concept of subjective normative reasons differs from the concept of justifiable reasons that I developed in Chapter 4. I explain why after presenting Richard Joyce's similar account of reasons.

6.2.2 Richard Joyce

Richard Joyce's example resembles that of Schroeder: Molly is thirsty. She believes, reasonably, that this cup contains water. It contains poison. Nevertheless, because she has good epistemic reasons for believing that the cup contains water, she has reason to drink from it (Joyce, 2001, pp. 53-55). Joyce defines objective and subjective reasons as follows:

> S has an objective reason to φ if and only if φ-ing will further S's ends.
>
> S has a subjective reason to φ if and only if S is justified in believing that she has an objective reason to φ. (Joyce, 2001, p. 53)[74]

When Joyce writes 'has … reason' here, he refers to reason relations, what agents have reason to do, not reason-explanans, that is, he is not claiming that the reason-explanans that gives the agent reason to act is the agent's belief.[75] If Joyce's analysis of reasons is expressed using an example, the distinction becomes clearer:

> Sally has a subjective reason to take the blue pills if and only if Sally is justified in believing that she has an objective reason to take the blue pills.

The reason-explanans for Sally to take the blue pills is missing from this account of reasons. Sally's reason-explanans could be that she loves the colour blue, that the blue pills will cure what ails her, that her doctor told her to take them, or many other things. The definitions that Joyce provides of these reason relations give no indication of whether Joyce takes the reason-explanans to be states of affairs, propositions or psychological states. At the time that Joyce wrote this book, he did not have a position on the metaphysical status of reason-explanans (Joyce, 2010, pers. comm., July).

---

[74] Joyce focuses on instrumentalism at this point, hence the emphasis on furthering the agent's ends.

[75] The distinction between the reasons that are taken to be reasons for action, reason-explanans, and the relation that exists when an agent has reason to act – reason relations – is explained in §2.3.

Although Joyce writes that an agent has a subjective reason to act in some way if and only if she is justified in believing that she has an objective reason to act in that way, Joyce presumably intends his analysis to be read as a counterfactual conditional; that is, Joyce would presumably agree with the following formulation of his concept:

> *S* has a subjective reason to *φ* if and only if *S* would be justified in believing that she has an objective reason to *φ*.

This slightly altered formulation allows an agent to have a subjective reason to act even when she does not currently have a belief about what she has reason to do. It also makes it clearer that there can be a third-person account of what an agent has reason to do that can agree with the agent's own analysis of what she has reason to do, or detect problems with the agent's analysis.

In a review of Richard Joyce's book, Jay Wallace rejects Joyce's claim that there can be subjective reasons (2003a). He means two things by this. First, there is not a kind of reason for action called a 'subjective normative reason' that is somehow different from an objective normative reason; all normative reasons for action are objective reasons. Second, the examples used to illustrate cases where people have subjective normative reasons for acting are not examples of agents having reason to act; that is, if my doctor tells me to take medicine that will harm me, I have no good reason to take the medicine. Given that on my account of justifiable reasons an agent may have reason to take medicine that will harm her if she would be justified in taking herself to have reason to take that medicine, Wallace's objection to Joyce's concept of subjective normative reasons appears to also be an objection to justifiable reasons.

When rejecting Richard Joyce's claim that agents with justified beliefs may have subjective reasons to act on those beliefs, Wallace writes,

> An 'objective' reason is a genuine normative reason … a consideration that really does count in favor of the person's [*φ*-ing]. 'Subjective' reasons are not really reasons at all, but rather beliefs of agents about what they have reason to do. (Wallace, 2003a)[76]

However, Wallace's objection does not meet its target. To see this, first consider the conception of reasons that Wallace uses here and the conception of reasons that he ascribes to Joyce, then contrast this understanding of reasons with Joyce's account.

Wallace describes an 'objective reason' as a consideration that favours an agent acting some way (2003a). Considerations that favour an agent acting in some way are reasons in the sense that they explain why the agent has reason to act in some way. Thus, Wallace takes the debate to be about the nature of what I call reason-explanans. Wallace then describes subjective reasons as 'beliefs of agents about what they have reason to do'. This association

---

[76] It is, perhaps, unfair to criticise Wallace on the basis of a minor comment made in the middle of a book review, but examining Wallace's criticism helps show the importance of distinguishing between reason relations and reason-explanans.

of reasons with beliefs is understandable, given that Joyce states that an agent has a subjective reason to act when she 'is justified in believing that she has an objective reason'. As I mentioned in §4.0, the use of the word 'belief' tends to strike immediate alarm bells in the minds of some philosophers working on practical reason. Because Wallace's focus is on the reasons that favour an action (the reason-explanans), rather than on an agent having a reason (the reason relation), Wallace assumes that Joyce is claiming that it is agents' beliefs that serve as reasons in the sense that they favour them acting in certain ways. He contrasts this with the objectivist position, where it is states of affairs, facts, or propositions that are reason-explanans for actions. But, this is a misreading of Joyce's position, albeit an understandable one.

As I wrote above, Joyce did not intend to make any claims about the ontological status of reason-explanans when he wrote *The Myth of Morality*. So, he was not claiming that reasons, as in considerations that count in favour of an agent acting in a certain way, are beliefs. When Joyce claims that an agent would be justified in taking herself to have reason to act in a certain way, he refers to a reason relation. It would be consistent with Joyce's claims to hold that the reasons that favour agents' actions are of a kind; that is, whatever the reason relation, the reasons that favour actions are states of affairs. Return to Joyce's example of Molly. She sits in her regular cafe feeling a little thirsty. Her regular pleasant waiter brings her a water jug filled with clear liquid. I take it that these states of affairs – Molly's thirst, the water-like liquid in the jug on her table and the way in which the water jug arrived at her table – are considerations that count in favour of Molly drinking from that jug. Assume that an invisible elf, peeved at Molly's failure to acknowledge his existence, turns the liquid in the jug to poison. It remains the case that Molly's thirst, the watery liquid in the jug on her table and the way in which the water jug arrived at her table favour her drinking from that jug.

Joyce was also not claiming that an agent needs to have a belief to have a subjective reason to act in some way, or that if she has a belief about what she has reason to do she has a subjective reason to act in some way. Whether Molly believes that she has reason to drink the water is irrelevant to whether it is appropriate for her to hold that she has reason to do so. Assume that Molly decided, based on no evidence at all, that invisible elves had poisoned her water. As a consequence of her batty reasoning, she forms the belief that she has reason not to drink the water. Molly's belief is true, but because Molly's belief is not justified, the fact that she believes she has no reason to drink from the jug does not give her a Joycean subjective reason not to drink from the jug. In spite of her beliefs, Molly has a Joycean subjective reason to drink from the jug, because if she carried out the reasoning that it is possible and appropriate for her to carry out, then that is the conclusion she would reach.

6.3 Justifiable reasons are not justifiable beliefs about objective normative reasons

Joyce's position is close to the one for which I argue; his analysis of subjective reasons describes a sufficient condition for an agent to be justified in taking herself to have reason to

act. So if Wallace's criticism of Joyce were successful, that would affect the plausibility of my claim that it is possible to give an account of normative reasons that apply to agents whose circumstances somehow prevent them from learning what they would have objective normative reason to do. However, Wallace's confusion about the reasons to which Joyce refers means that his criticisms have no impact on the normativity of Joyce's concept of subjective reasons and no impact on the concept of justifiable reason relations that I developed in Chapter 4. I have different grounds for rejecting Joyce's account, along with that of Cullity and Gaut: justifiable reasons are not beliefs about objective reasons, they are not even justified beliefs about objective normative reasons.

Cullity and Gaut, and Joyce, appear to intend to describe a concept of reasons that corresponds to my account of justifiable reasons, and the concepts of subjective normative reasons that they develop are normative. They ensure that their accounts of subjective normative reasons are normative by including a requirement that for an agent to have a subjective normative reason to act, the agent's beliefs must be justified. The difference between my account of justifiable reasons and Cullity and Gaut's, and Joyce's accounts of subjective normative reasons arises because they define subjective normative reasons in terms of objective normative reasons. Because they take this approach, their accounts of reasons are unable to explain the reasons that real human agents have for acting.

There are two reasons for rejecting the claim that the normative reasons that real agents have for acting are justified beliefs about objective reasons. The first difference between the accounts arises because of the use of the word 'belief'. Justifiable reasons and reason relations are also not beliefs about what an agent has objective normative reason to do.[77] They are not even justified or justifiable beliefs about what an agent has objective normative reason to do. First, the existence of justifiable reasons and reason relations is independent of anyone's beliefs about what they have reason to do. To describe justifiable reasons as 'justifiable beliefs about objective reasons' invites the mistaken conclusion that it is the beliefs that determine what someone has a justifiable reason to do. If justifiable reasons were to be defined in terms of objective normative reasons, the definition of such reasons would need to be something like:

> *A* has a justifiable reason to *φ* in *C* if and only if, in *C*, she would be justified
> in believing that she has an objective reason to *φ*.

I think that Cullity and Gaut, and Joyce, intended to claim something like this in their accounts of subjective normative reasons.

Using beliefs to explain reason relations is unwise, even if the fact that an agent would be justified in believing she had an objective reason to act in some way were necessary and sufficient for that agent to have a justifiable reason to act. As soon as psychological states are given a key role in explicating notions of reasons, the facts and reasoning that lie behind those beliefs are de-emphasised, and the psychological states themselves become the focus.

---

[77] This is discussed in Chapters 4 and 8.

This change in focus seems to occur even when the psychological states are embedded within counterfactual conditionals. This is presumably part of the reason for the misunderstanding that led Wallace to reject the concept of subjective normative reasons. However, on its own, this point only shows that it is important for people who give psychological states a key role in explicating notions of reasons to tightly define the role they give psychological states.

The second difference between my account of justifiable reasons and the accounts of subjective normative reasons given by Cullity and Gaut, and Joyce arises because these theorists explain subjective normative reasons in terms of objective reasons. As I argued in §4.4, justifiable reasons cannot be explained in terms of objective normative reasons in this way.

If Cullity and Gaut, and Joyce had claimed only that they were describing sufficient conditions for an agent to have a subjective normative, or justifiable, reason to act, they would have been correct. If an agent's situation were such that she would be justified in believing that she had an objective reason to act in some way, then she would have a justifiable reason to act in that way. To see this, imagine that everything about your situation and your doctor's words leads to one conclusion: the medicine will cure you. You form a belief that you have reason to take the medicine, because you believe that your understanding of the state of the world is correct, your reasoning is correct and your understanding of the consequences of taking the medicine is correct. And you are justified in forming this belief. It is highly unlikely that you think 'I believe I have an objective reason to take the medicine', but, arguably, your certainty about your understanding of your situation, your reasoning and the consequences of your taking the medicine amounts to you having such a belief. In this situation, Cullity and Gaut, and Joyce would claim that you have a subjective normative reason to act, and I would accept that you have a justifiable reason to act. But an agent can have a justifiable reason to act when it is not the case that she would be justified in believing that she has an objective reason to act.

It is not a necessary condition for an agent to have a justifiable reason to act that she would be justified in believing that she has an objective reason to act. In situations of uncertainty agents may not be justified in forming any belief about what they have objective reason to do. Imagine that the information you have is equivocal, but that if you reasoned to the best of your ability you would conclude that taking the medicine is a good bet. If you think that taking the medicine is a good bet, but not a sure thing, you do not have grounds for believing that you have an objective reason to take the medicine. You know that you do not have all the relevant information about your illness and the medicine, so however perfect you think your reasoning is, you would not be justified in concluding that you have an objective reason to take the medicine. However, if you think that taking the medicine is a good bet, you would be justified in taking yourself to have reason to take the medicine even though you do not believe that you have an objective reason to act, that is, even though you realise that your taking the medicine might not have the effect that you think it will have. Thus, the account of subjective normative reasons given by Cullity and Gaut, and Joyce does not correspond to my account of justifiable reasons.

In Chapter 4, I argued that justifiable reasons are not a form of motivating or objective normative reasons. The arguments that I have given show that justifiable reasons are not a form of subjective normative reasons, but also reiterate my arguments against one approach to trying to explain justifiable reasons in terms of objective reasons.

6.4 Subjective normative internalist reasons

Peter Railton sketches an account of what he calls 'subjective reasons' that comes close to being an account of what I would call an 'internalist account of justifiable reasons' (2008). Internalists about reasons hold, roughly, that for an agent to have reason to act in some way, the agent must potentially be able to be motivated to act in that way.[78] Like the forms of subjective reasons already discussed in this chapter, aspects of Railton's account of subjective reasons conflict with my account of justifiable reasons. However, the similarities are suggestive.

Like my distinction between justifiable reasons and objective normative reasons, Railton's account of what he calls 'subjective' and 'objective' reasons is based on distinguishing between levels of idealisation. Both forms of reason are normative, but one form results from a higher degree of idealisation about agents' knowledge and reasoning. Railton's account of subjective and objective reasons is as follows:

> Subjective reasons support directive judgments of what an agent ought to do or would be subject to rational criticism for failing to do, so they must lie within the agent's epistemic ken, and must have the power to move him to thought and action more or less as he is; otherwise, we cannot expect him to see the rationale of the ought or hold him responsible for failing to feel its action-guiding force. Objective reasons also have an internal connection to the agent's motivational structure, not as it actually is, but as it would be under an idealization of his information, imagination, experience, acumen and understanding. Objective reasons support evaluative judgments of what would be best for the agent … (Railton, 2008, p. 238)

Railton's account of both subjective and objective reasons is internalist because what agents can have subjective and objective reason to do is limited to actions that the agents could be motivated to carry out. The similarity between my account and Railton's is immediately apparent. Like me, he makes an agent's epistemic circumstances key for determining what the agent has reason to do. And, like me, he assumes that there can be two forms of normative reason.

There are three differences between Railton's briefly sketched account of subjective reasons and my account of justifiable reasons. Two of those are minor. The third is the subject of the next chapter.

---

[78] Internalism and externalism about reasons are discussed in more detail in Chapter 7.

First, Railton does not specify that subjective reasons are reasons for action that agents would be justified in taking to be reasons for action. However, Railton's comments about agents being 'subject to rational criticism' if they fail to see the force of such reasons seems to indicate that he takes justification to be key to an agent having what he calls 'subjective reasons' for acting. Presumably an agent's rationality would be criticised if she acted for reasons that she was not justified in taking to be reasons. Making this explicit would, again, require only a minor alteration to Railton's account.

Second, Railton must be referring to overall reasons rather than pro tanto reasons, because agents would often be praised rather than criticised for failing to act on pro tanto normative reasons. This seems an insignificant difference between my account of justifiable reasons and Railton's account of subjective reasons, but if Railton is referring to overall reasons, his claim that agents ought to act on reasons is still stronger than the claims I argued for in §2.2. For example, it seems conceivable that an agent might not be criticised for failing to act on a supererogatory reason, even if it is what she has overall reason to do. Like the other differences mentioned, it would be easy to alter this aspect of Railton's account of subjective reasons so that it says something like 'Overall subjective reasons support directive judgments of what an agent ought to take herself to have some reason to do'.

All things considered, it seems that Railton's account would only need to be altered slightly to be an account of what I would call 'justifiable internal reasons', that is, justifiable reasons that only exist when they are appropriately related to agents' motivational sets. This is, however, not an indication that justifiable reasons must be internal reasons. In spite of first impressions, the debate about internalism and externalism about reasons is orthogonal to the debate about the existence of what I have been calling 'justifiable' and 'objective normative reasons'. Justifiable reasons are not a form of internal reason, nor does the existence of justifiable reasons entail that objective normative reasons are internal reasons. This is the fourth and most significant difference between my position and that of Railton. I argue for this claim in the next chapter.

6.4 Conclusion

Human agents have limited knowledge and reasoning abilities, yet they constantly make decisions about how they will act. Real agents think and speak about 'reasons for acting' and what they 'have reason to do' even when they know that they do not have all the relevant information or every useful reasoning skill. This chapter described three approaches to understanding the reasons agents have when their knowledge does not correspond to that of ideal agents. I argued that the concepts of subjective normative reasons described by Sie, Slors and van den Brink, Schroeder, Cullity and Gaut, and Joyce do not give a correct account of real agents' reasons for acting.

The concepts of subjective normative reasons described by Maureen Sie, Marc Slors and Bert van den Brink, Mark Schroeder, Garrett Cullity and Berys Gaut, and Richard Joyce fail to

adequately capture the notion of common-sense normative reasons for action for which I argue (Cullity & Gaut, 1997; Joyce, 2001; Schroeder, 2008c, 2009; Sie et al., 2004). In their development of subjective normative reasons, they correctly claim that you can have reasons to act that are based on a partial, and potentially misleading, understanding of the world. However, the approach they take leads philosophers to either conclude that the reasons that matter are objective reasons and almost ignore real agents' reasons or conclude that real agents' reasons for acting are not normative reasons.

The concepts of subjective normative reasons introduced by Sie, Slors and van den Brink and Schroeder seem intended to be normative, but justification is not given the role it needs to have for these to be normative reasons for acting. At best, Sie, Slors and van den Brink introduce a concept of motivating reasons that just happen to be normative. The supposed 'normativity' of these reasons is doubtful, as there is nothing in Sie, Slors et al.'s account that makes it anything other than a happy coincidence that their subjective normative reasons are normative. Other differences aside, justifiable reasons are not a form of motivating reason. Schroeder's account of 'subjective normative reasons' is also radically different from my account of justifiable reasons. Like Sie, Slors, et al.'s account, it is doubtful that Schroeder's subjective normative reasons are normative. Although Schroeder's comments and his use of examples suggests that he intends his account of subjective normative reasons to be normative, his definition of such reasons suggests that agents can have a Schroederian subjective normative reason to act even when their belief that they have reason to carry out the action is unjustified. Schroeder's definition of 'subjective normative reasons' also conflicts with my account of justifiable reasons because agents can have a justifiable reason to act when they have a justifiable, but false, belief about the appropriateness of their goal. The way in which Schroeder relates subjective normative reasons to objective normative reasons means that if an agent has a mistaken belief about her goal, she has no subjective normative reason to act to bring about that goal.

Cullity and Gaut, and Joyce introduce genuinely normative concepts of reasons, but they claim that an agent has a subjective normative reason to act, when she has a justifiable belief that she has an objective reason to act some way. I argued that this fails to recognise that an agent can have a justifiable reason for acting in a certain way even when she would not be justified in believing that she has an objective reason to act in that way.

Peter Railton sketches an account of subjective reasons that, in spite of its differences from my account of justifiable reasons, comes closest to capturing my position (Railton, 2008). However, Railton's account appears to assume that his forms of subjective and objective reasons must also be internal reasons. In the next chapter I show that the distinction between justifiable reasons and objective normative reasons is not the same as the distinction between internal and external reasons. I also argue that the notion of justifiable reasons is consistent with Bernard Williams' arguments in support of internalism.

7. Internalism and Justifiable Reasons

7.0 Introduction

When I first began to search for a theory of normative reasons that was responsive to the limitations of human agents, internalism about reasons seemed the most plausible option, and I have previously defended that position (Mason, 2006). Internalists about reasons tie reasons to actual agents by arguing that an agent only has a normative reason to act some way if she would be motivated to act in that way after correct deliberation.[79] The competing account of reasons – externalism – states that agents sometimes have reason to carry out actions that they could never be motivated to perform.[80] Peter Railton also seems to see compatibility between internal reasons and something like my account of justifiable reasons. Thus at the end of the last chapter, I described Railton's account of what he calls 'subjective reasons', that is, reasons that 'lie within the agent's epistemic ken', and I claimed that Railton's brief outline of an account of subjective reasons was close to being an internalist account of justifiable reasons (2008, p. 238). However, in this chapter I argue that there is no necessary connection between justifiable reasons and internalism about reasons; the concept of reasons that follows from internalism is significantly different from my concept of justifiable reasons, and external reasons differ significantly from the forms of objective normative reasons discussed so far. My account of justifiable reasons neither presupposes, nor supports, internalism about reasons, nor are justifiable reasons internal reasons.

Not only is the concept of internal reasons significantly different from that of justifiable reasons, some arguments in support of internal reasons seem also to be arguments against the normativity of justifiable reasons. Bernard Williams and Michael Smith argue that what an agent has normative reason to do depends on facts about the agent's situation (Smith, 1994, 1995a; Williams, 1981a, 1995a, 1995b, 2001). If, for example, you want a gin and tonic and there is petrol in the gin bottle, they claim that you have no reason to drink from the bottle. Williams and Smith argue that this dependence on truth is a necessary requirement for their accounts of internal reasons to be accounts of normative reasons. This means that their arguments in support of internalism appear to be directed against accounts of reasons that resemble my account of justifiable reasons.

I do not accept that Williams' and Smith's arguments show that my account of justifiable reasons is not an account of normative reasons. Instead, I argue that there are grounds for believing that Williams' commitment to internalism about reasons was based partly on the same grounds that lead me to argue in favour of the concept of justifiable reasons. I argue that Williams would have achieved many of his ends if he had adopted my notion of justifiable

---

[79] See, for example: (Darwall, 1997; Finlay, 2006; Korsgaard, 1986; Smith, 1994, 1995a, 1995b; Williams, 1981a, 1995a, 1995b)

[80] See, for example: (Dancy, 2000; Finlay, 2006; Scanlon, 1998; R. J. Wallace, 2006)

reasons, and some of his own presuppositions sit more easily with my position on justifiable reasons than his theory of internal reasons. So, the similarities between some accounts of internal reasons and justifiable reasons are more than just misleading surface appearances, even though internal reasons and justifiable reasons differ.

In §7.1, I argue that in spite of the obvious similarities between internal reasons, subjective reasons and justifiable reasons, and between external reasons and objective normative reasons, the distinction between internal and external reasons differs from the subjective–objective reasons distinction, and the justifiable reasons distinction is different again. In §7.1.1 I explain the basic differences between these distinctions. In §7.1.2, I argue that there can be internalist accounts of objective normative reasons. Finally, in §7.1.3, I argue that justifiable reasons need not be internal reasons.

Williams' and Smith's arguments about internalism emphasise the importance of correcting false beliefs, correcting faulty reasoning, and learning relevant true beliefs to determine what an agent has reason to do. In §7.2, I argue that these arguments do not undermine my account of justifiable reasons. Williams gives several arguments in support of his claim that correcting false beliefs and faulty reasoning, and learning relevant true beliefs is necessary for learning what agents have normative reason to do. He claims that correcting false beliefs and faulty reasoning, and learning relevant true beliefs makes normative reasons appropriately normative, appropriately connects reasons to the agent's rationality, corresponds to requirements for beliefs and reasoning held by agents themselves, and, ensures that informed agents will agree about what there is reason for a particular agent to do. I look at each of these claims in turn, and argue that even though justifiable reasons only involve correcting beliefs and reasoning in ways that are appropriate for agents, the normativity of justifiable reasons is not undermined by any of these claims.

In the last section of this chapter, §7.3, I show that although he appears to argue in ways that conflict with justifiable reasons, Bernard Williams' approach to reasons and ethics is compatible with the idea that justifiable reasons are an appropriate concept for explaining real human beings' reasons for action.

7.1 Internal–external, subjective–objective, and justifiable reasons distinctions

At first sight, there may appear to be a marked similarity between internal reasons and subjective or justifiable reasons, and a similarity between external reasons and objective normative reasons. An agent's potential psychological states, specifically, her capacity to be motivated to act in certain ways, affect what that agent has internal reason to do. What an agent can have justifiable reason to do is also limited by that agent's potential psychological states, because an agent only has justifiable reasons for acting if it is possible and appropriate for the agent to take herself to have such reasons for acting. Moreover, as argued in §4.3, limits on what can motivate an agent affect what the agent has justifiable reason to do, just as

they limit what the agent has internal reason to do. In apparent contrast, external reasons and objective normative reasons can exist independently of agents' motivational states. Externalists about reasons hold that there can be reasons that are not connected to motivation in any of the ways that internalists claim; they may, for example, argue that normative reasons arise from facts about the world not from facts about agents' hypothetical or actual motivational states. This position seems, on the surface, similar to the objective normative reasons theorists' claim that it is usually facts that determine what agents have reason to do. In spite of these similarities, the justifiable, subjective–objective and internal–external reasons distinctions differ.

7.1.1 Internal–external, subjective–objective, and justifiable reasons focus on different elements of the reason relation

Internalists, externalists, objectivists, subjectivists and I all make claims or assumptions about the nature of the reason relation. However, debates between internalists and externalists tend to assume that correct accounts of normative reasons are accounts of objective normative reasons. In contrast, accounts of subjective reasons and my account of justifiable reasons tend to qualify reasons in a way that brings them into contrast with accounts of objective normative reasons. The analysis of reason relations explained in §3.2 can be used to show these differences:

STATE	A state of affairs or proposition.	The person collapses.
END	Given the state of affairs, an action will yield an effect.	The collapse is a sign that helping the person might improve his functioning.
VAL	The effect has positive value.	Improving the collapsed person's functioning would have some positive value.
SIG	The positive value of the action remains significant in spite of any negative values produced by the action.	Helping the collapsed person would realise that value without conflicting with other values to such a degree that it would be implausible to say there is reason to do it.

Arguments about the correct account of objective reasons tend to be arguments about the correct account of *STATE, END, VAL* and *SIG*. In other words, if we had an appropriate idealised understanding of the world and (human) agency, what kinds of things could *STATE, END, VAL* and *SIG* be? As we have seen, accounts of objective normative reasons tend to claim that *STATE* and *END* are states of affairs or true propositions, not beliefs.

Aspect of the reason relation		Objective	Subjective
STATE	A state of affairs.	Reasons are states of affairs or propositions.	Reasons are beliefs or other information holding mental states.
END	Given the state of affairs, an action will yield an effect.		
VAL	The effect has positive value.	Reasons are states of affairs or propositions.	Reasons are motivations, desires or other pro-attitudes.
SIG	The positive value of the action remains significant in spite of any negative values produced by the action.		

As will be seen in the next chapter, of objective normative reasons theorists take a range of things to be potentially capable of being *VAL* and *SIG*, some of which might be considered 'subjective' in the sense indicated in the table above, and some of which might be considered 'objective'. Debates about the internal-external reasons distinction tend to also focus on the kinds of things that play the roles of *VAL* and *SIG* within reason relations, that is, what kinds of things can count as values within the reason relation.[81] Externalists claim that there are at least some reasons, usually moral or prudential reasons, where *VAL* and *SIG* are objective. Internalists about reasons are subjectivists about *VAL* and *SIG*; that is, they hold that agents only have reason to act in some way when *VAL* and *SIG* are appropriately related to the agents' pro-attitudes. Kantian internalists, such as Christine Korsgaard and Michael Smith, refine this position by arguing that by placing the correct requirements on agents' desires, you can get an account of objective values – at least in the intersubjective sense of 'objective' (Korsgaard, 1986; Smith, 1994). This is discussed in more detail below. So, there is nothing about internal reasons as such that makes them subjective reasons rather than objective reasons.

The connection between motivational sets and the value of acting in certain ways can be seen in Williams' and Smith's forms of internalism. It is also revealed by the way in which Smith rejects some aspects of Williams' position while developing other aspects.

On Williams' preferred account of internal reasons, '[An agent] *A* has a reason to *φ* only if there is a *sound deliberative route* from *A*'s subjective motivational set [labelled "*S*"] to *A*'s *φ*-ing' (2001, p. 91, italics in original). According to Williams, a person's subjective motivational set includes more than is usually meant by the term 'desire'. A person's subjective motivational set 'can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects … embodying commitments of the agent' (1981a, p. 105). These motivators can be altruistic, egoistic, or neither altruistic nor egoistic. What someone has reason to do is limited by her subjective motivational set, but

---

[81] See §3.2 for an explanation of what is meant by 'value' here.

is not limited to what she is currently motivated to do. This is part of what distinguishes Williams' internal reasons from motivating reasons. Instead, what an agent has reason to do is determined by what the agent would be at least somewhat motivated to do were she to carry out sound deliberation about her situation. This deliberation may lead an agent to develop new desires or values and discard existing elements of her motivational set.

Michael Smith develops an internalist theory of reasons based partly on his rejection of Williams' semi-Humean approach to normative reasons, but his approach is similarly focussed on values and desires. According to Smith, for an agent to identify her normative reasons for action, she 'must have no [relevant] false beliefs … all relevant true beliefs … [and] must deliberate correctly' (1994, p. 156; 1995a, p. 112).[82] On Smith's account, correct deliberation requires that an agent's motivational set is systematically justifiable; that is, that agents' 'underived desires [must] form a maximally coherent and unified desire set' (1994, pp. 156-161; 1995a, pp. 114-117). To make their desires coherent, Smith suggests that agents think of broad general desires that unify their particular desires. He claims that this process of forming a coherent and unified set of desires will involve discarding desires that conflict with agents' broad general desires and gaining new desires that do cohere. Thus, on Smith's account, desires can be formed through deliberation. It is the desires that we would keep if we were rational that are the basis for normative reasons.

Smith argues that if agents have systematically justified, hence, maximally coherent and unified desire sets, agents' beliefs about what there is reason for any agent to do in a particular situation will converge and every agent would agree on what an agent has reason to do in some set of circumstances (1994, pp. 164-174; 1995a, pp. 117-125; 1995b, pp. 294-296). For this reason, Smith rejects what he describes as the 'relativist' aspect of Williams' work, that is, the way that Williams' account entails that agents can have differing reasons for action in the same external circumstances. This supposed relativism arises because of Williams' claim that rational deliberation has a limited ability to alter agents' existing motivational sets. In contrast, Smith claims that it is a conceptual truth that normative reasons are non-relative. He claims that, if they reason appropriately, people will agree about what kinds of outcomes are worth bringing about, where 'outcomes' is to be read in the broadest sense of the word, that is, as including the form of actions, not just the consequences of actions. (If Smith is right, this unified understanding of what is of value is tremendously important for moral philosophy.)

Unlike the internalist, externalist and objectivist theories discussed so far, my theory of justifiable reasons is concerned more with when *STATE, END, VAL*, and *SIG* serve as normative reasons for real human agents than with whether these aspects of the reason relation are objective or subjective. I claim that an agent has a justifiable reason for acting when the following criteria are satisfied:

---

[82] The word 'relevant' is my insertion, to increase the plausibility of Smith's position; that is, I assume that Smith does not hold that to determine what an agent has reason to do we must assume that he has no false beliefs of any kind.

> (*STATE*) There is a state of affairs that it is appropriate for the agent to be aware of.

And, it is appropriate for the agent to hold that:

> (*END*) the state of affairs is such that if she acts in a certain way, it will have a certain effect;
>
> (*VAL*) the result of acting in that way will have a certain value; and
>
> (*SIG*) the value of acting in that way is not outweighed by any negative effects of acting in that way.

In §5.2.1, I argued that it is possible to give an account of justifiable reasons where *STATE* is taken to be a state of affairs rather than a belief about a state of affairs. I haven't taken a position on the nature of *VAL* and *SIG*. When I explained what I meant by *VAL* and *SIG* in §3.2, I tried to allow for as many ways of analysing *VAL* and *SIG* as possible. I avoided making claims that would commit me to any position on debates about the relationship between reasons, desires, value and motivation, including debates about the Humean Theory of Reasons.

I have argued that those involved in the externalist–internalist debate typically assume that reasons are objective, and that the issue of subjective and justifiable reasons is different again. Since the externalist–internalist distinction is orthogonal to my account of justifiable reasons, there can be internalist or externalist accounts of justifiable reasons. To clarify this I show that there can be internalist accounts of objective normative reasons and externalist accounts of justifiable reasons.

7.1.2 The internalist's reasons can be objective normative reasons

Although Railton's account of subjective reasons described at the end of Chapter 6 seems to be an internalist account of justifiable reasons, it is possible to be an internalist about normative reasons without accepting my account of justifiable normative reasons (Railton, 2008). There is nothing about internalism that would commit an internalist to accepting the claim that what an agent has reason to do is limited by what the agent can learn she has reason to do. Williams' internal reasons are a variety of objective normative reason. To see this, consider Williams' use of the example of the gin and tonic drinker. Williams asks us to consider whether an agent who wants a gin and tonic has reason to pour the contents of the gin bottle in front of him into a glass, add some tonic, and drink. The catch is that this gin bottle contains petrol (Williams, 1981a, p. 102). Williams argues that an agent does not have reason to act in some way if his motivation is based on a false belief or if he would not be motivated to act in that way if he knew all the information relevant to acting in that way (1981a, p. 103). So, to determine what the man has reason to do, we should correct his false belief about the contents of the bottle. After doing this, we can conclude that he has no reason to pour the stuff in the bottle into a glass and drink it. This requirement on normative reasons makes Williams' account of internal reasons an account of objective normative reasons because what the agent has reason to do is determined using states of affairs that the agent

may not be able to access or reasoning that the agent may not be able to carry out. Although I consider an alternative reading of Williams' claims in §7.3 below, on what seems a reasonably standard reading, his position differs markedly from the requirements I give for an agent to have a justifiable reason for acting.

Michael Smith's account of internal reasons is also an account of objective normative reasons. According to Michael Smith, to identify what an agent has normative reason to do, we must correct any false beliefs she has, add in any true beliefs that are relevant to the action and deliberate correctly on the facts in a way that produces a coherent set of beliefs and desires (1994, p. 156; 1995a, p. 112). Although Smith refers to agents' beliefs, those beliefs must be corrected by an all-knowing observer. What an agent has reason to do is determined by the facts of her situation, which makes Smith's account of reasons an account of objective normative reasons.

David Sobel argues that Williams' internalism could be treated as either an objective or subjective theory of reasons (2001, pp. 473-474). Sobel calls his objectivist reading of Williams' internalism 'tracking internalism'. On a tracking internalism account of reasons, what agents can be motivated to do after sound deliberation reliably tracks what agents have reason to do. This approach takes objective reasons to exist independently of agents, rather than sound deliberation making it the case that the agent has such a reason. Sobel does not think that Williams intends to develop an account of tracking internalism, and this seems correct. On Williams' account, sound deliberation and the contents of the motivational set determine what the agent has reason to do rather than track what the agent has reason to do. Sobel concludes that Williams' internalism is subjectivist.

Sobel's argument does not show that Williams' account of reasons is not an account of objective normative reasons as I use the term. The range of ways in which the term 'objective normative reason' is used means that it is important to check for ambiguities when interpreting claims about whether a theory of reasons is objective or subjective. 'Objective normative reason' as I use the term, does not exclude states of affairs that exist as reasons only when agents are able to determine that those states of affairs give them reason to act in some way. There is nothing peculiar about taking this approach to objective normative reasons. Consider Dancy's position on reasons (2000, 2003, 2009). Dancy takes himself to have developed an externalist account of objective normative reasons. But Dancy claims that any theory of reasons must allow for an epistemic filter that prevents states of affairs that agents cannot access from serving as reasons for those agents; an agent's ability to become aware of a state of affairs is what enables that state of affairs to be a reason (2000, pp. 56-59, 65-66). The fact that Dancy's theory would be subjectivist on Sobel's version of the subjective–objective distinction shows that Sobel is using the term differently from the way it is used here.

Several other philosophers argue that Bernard Williams provides an objectivist account of internal reasons. Kieran Setiya, Mark Schroeder and Peter Railton all point to Williams' use of his petrol and tonic example to support this claim (Railton, 2008, p. 238; Schroeder, 2008c,

p. 13, fn 18; Setiya, 2004, p. 271).[83] Setiya, Schroeder, and Railton argue that what makes Williams' internalist theory of reasons objective is that it is grounded in actual states of affairs rather than beliefs about states of affairs. This understanding of 'objective normative reasons' links the objectivity of such reasons to *STATE* and *END* in a way that corresponds to my arguments; it supports the currency of my position on the distinction between subjective and objective reasons as well as my claim that Williams' internal reasons are objective normative reasons in my sense of the term.

7.1.3 Justifiable reasons and the internal–external distinction

Although I claim that my account of justifiable reasons is not committed to any particular position on internalism and externalism about reasons, two aspects of my position may suggest that I give an account of internal justifiable reasons. First, my approach to value might be thought internalist because of the restriction I place on when any possible content for *VAL* and *SIG* can form the basis for reasons for action. Second, my argument in §4.3 that agents' motivational systems can affect what they have reason to do resembles internalists' claims about the relationship between motivation and reasons.

Externalists argue that agents can have moral or prudential reasons to act in certain ways whether or not they could be motivated to act in those ways. Intuitively, we seem to want to say that people have reason to act to realise certain values, even when they could never be brought to accept that they have reason to act in such ways. Daily life and the popular media give us plenty of examples of people whose motivational sets seem to contain mistaken dispositions. It is easy to make sense of the claim that a chronic manipulator has reason to treat her partner with respect even when she has nothing in her motivational set that would lead her to alter her behaviour. It also seems to make sense to claim that everyone has reason to help the homeless when they can, even those who could never be brought to see they have reason to do so. Examples such as these work best with moral reasons, because people are inclined to hold that others have moral reasons to act in certain ways irrespective of their motivational capacities or inclinations. However, the same intuition pump can be used for prudential reasons. If someone needs to get somewhere fast, he seems to have at least a pro tanto reason to take the quickest route. He seems to have reason to do this even if it would not be possible to motivate him to take that route because he has an insurmountable fear of taking the direct route. The externalist response to examples such as these is to say that there can be agent-independent normative reasons for people to act in certain ways.

An externalist about justifiable reasons could argue that it can be appropriate for an agent to realise that she has reason to act some way even though she would not be motivated to act in that way. The simplest way to argue for externalism about justifiable reasons is to claim that it is possible for agents to identify reasons without being motivated to act on those reasons.

---

[83] Schroeder writes that Williams seems to have believed that his internal reasons account works for both objective and subjective accounts of reasons, but Schroeder uses the term 'subjective reasons' in a unique way (Schroeder, 2008b; 2008c, p. 13, fn 18; Williams, 1981a, pp. 102-105; 2001, pp. 92-93).

Imagine, for example, a psychopathic utilitarian; such a person might accept that everybody has reason to act to maximise expected happiness, but feel no inclination at all to do so. He has justifiable reasons for carrying out certain actions, but nothing in his motivational set inclines him to act as he claims he has reason to act. He seems, therefore, to have external justifiable reasons for acting.

This argument that someone could be an externalist about justifiable reasons is not an argument in support of being an externalist about justifiable reasons. The argument about the possibility of psychopathic utilitarians is purely theoretical. I have no idea whether someone who carried out practically appropriate reasoning and actions could be a psychopathic utilitarian. An internalist about justifiable reasons could also argue that for an agent to have a justifiable reason to do something, he must not just be able to realise that he has reason to do it, he must also be able to bring himself to do it.[84] And, arguably, if he can't be motivated to do it, there is a sense in which he can't do it.

My account of justifiable reasons puts a constraint upon *VAL* and *SIG*; the content of *VAL* and *SIG* must be limited to estimates of value that it is appropriate for agents to make. This restriction excludes certain kinds of estimates of value. Some utilitarians argue that the value of an action is given by the value of the actual outcome of the action (Shaw, 1995, pp. 115-116; Tännsjö, 2010, pp. 32, 34).[85] Agents are commonly unable to come up with accurate estimates of the values of the consequences of potential actions, so this claim is incompatible with my account of justifiable reasons. This does not show that justifiable reasons are internal reasons because not all externalists are actual value utilitarians. Jonathan Dancy and Thomas Scanlon both argue for externalism without committing themselves to any form of consequentialism (Dancy, 2000; Scanlon, 1998). One could also presumably be an expected value consequentialist and ethical intuitionist, claiming that agents can perceive the value in features of the world while arguing in support of reasons externalism. And, such a position might be developed in a way that was consistent with my position on justifiable reasons.

The relationship between justifiable reasons and motivation may seem to inextricably link justifiable reasons to internalism. In §4.3, I argue that an agent's motivational system can affect what the agent has justifiable reason to do. I describe experiments that suggest that features of an agent's situation can affect the agent's motivation in ways that will affect what the agent is able to perceive about the world, and, thus, affect what the agent is capable of doing (Fine, 2010, pp. 20-21; Ickes, 2003; Ickes et al., 2000; Klein & Hodges, 2001). One example of this is the way that a two dollar reward for correctly interpreting another person's thoughts and feelings improved the accuracy of agents' predictions of another person's thoughts and feelings (Klein & Hodges, 2001, p. 727). Another example is the way in which women's empathic ability improved when they were reminded that they were women and, as such, are expected to be good at empathising (Ickes, 2003, p. 135). Consider the results of

---

[84] See, for example, Railton's account of subjectivist internalism about reasons (2008).

[85] The relationship between objective normative reasons, justifiable reasons and moral reasons is discussed in Chapter 8.

this last series of experiments. It seems unlikely that most women know that their motivation to accurately interpret other people's thoughts and feelings increases when they call to mind that they are women and, therefore, are expected to be good at the task. This ignorance of the effect of circumstances on motivation explains why the results of these experiments were so surprising. It would not be appropriate for agents to constantly check scientific literature for new findings on motivation, conduct experiments to investigate the effects of various background conditions on their motivations, or wait for scientists to conduct the relevant experiments before they act. It follows that women cannot usually control this aspect of their ability to empathise. Other people's thoughts and feelings often give agents reasons to act. Therefore, the results of this experiment suggest that agents' motivational systems sometimes limit agents' abilities to access information about other people that will affect what they have reason to do. So, what agents have justifiable reason to do may be limited by the way in which features of agents' situations affect their motivation. This shows that agents' motivational systems sometimes affect what agents have justifiable reason to do.

The similarity between justifiable reasons and internalist accounts of reasons does not mean that my account of justifiable reasons is an internalist account of reasons, at least not in the usual sense of that term. The relationship between agents' motivational sets and what agents have justifiable reason to do differs from the role of motivational sets in internalist theories of reasons. Internalists hold that an agent has normative reason to act in some way only if she would be motivated to act in that way were she to deliberate correctly about the facts of her situation and the contents of her motivational set (Smith, 1994, 1995a; Williams, 1981a, 1995a, 2001). In contrast, on my account of justifiable reasons, agents' motivational attributes are only relevant when they affect agents' access to information about the world. The theory of justifiable reasons that I have developed is not based on an assumption that people must be capable of being motivated to act in a certain way for them to have reason to act in that way.

In summary, there is no necessary association between objective normative reasons and external reasons, nor between justifiable reasons and internal reasons. Although arguments about internal and external reasons, arguments about subjective and objective reasons, and arguments about justifiable reasons are all aimed at uncovering the nature of reason relations, these arguments are about fundamentally different issues. In the next section, I show that the distinction between justifiable reasons and internal reasons is also apparent in the arguments that Williams and Smith give to show that internal reasons are normative.

7.2 Williams' and Smith's arguments that what an agent has reason to do is a product of factually and rationally correct beliefs do not undermine justifiable reasons[86]

Unlike my account of justifiable reasons, Williams and Smith claim that what an agent has internal reason to do is derived from idealised knowledge and reasoning that will sometimes

---

[86] This section is taken, with minor revisions, from an earlier paper (Mason, 2006).

go beyond the capabilities of human agents. Williams gives four interrelated reasons for believing that correcting false beliefs and learning relevant true beliefs is necessary for learning what agents have normative reason to do, and variations of two of these arguments are also given by Smith. However, these arguments do not show that my concept of justifiable reasons is inconsistent; in particular, they do not show that justifiable reasons are not a form of normative reason.

7.2.1 Correcting false beliefs and faulty reasoning and learning relevant true beliefs makes normative reasons appropriately normative

Williams writes that for 'a statement of the form "*A* has reason to *φ*" … [to have] *normative force* …. [it must] go beyond what that agent is already motivated to do – that is, go beyond his already being motivated to *φ*' (1995a, p. 36, italics in original). Correcting facts and reasoning is one way in which the internal reasons conception of '*A* has reason to *φ*' described by Williams and Smith goes beyond what an agent believes she has reason to do before she deliberates.

Assume, however, that Williams' and Smith's requirements on deliberation are weakened, so that what agents have reason to do is limited to what it is possible and appropriate for those agents to judge they have reason to do. An agent must still correct false information that it is possible and appropriate for her to correct and similarly search out relevant missing material, as well as deliberating in other ways when required. Establishing what an agent has justifiable reason to do still requires that agents' motivations, beliefs and reasoning are corrected to the extent that agents are capable of correcting them. Thus, the limitations that I place on the degree of idealisation necessary for an agent to have justifiable reason to act in some way does not prevent justifiable reasons from being normative reasons.

7.2.2 Correcting false beliefs and faulty reasoning and learning relevant true beliefs appropriately connects reasons to the agent's rationality

Williams argues that to say that *A* has a reason to *φ* when he is not 'factually and rationally correctly informed' would 'imply in effect that the internal reason conception is only concerned with explanation, and not at all with the agent's rationality' (1981a, pp. 102-103; 1995a, p. 37). It is always possible to give an internal reasons account of an agent's reasons for actually acting, because whenever an agent acts, it is possible to give an explanation of the action that refers to the agent's motives. But an agent can act in some way without there being normative reason to act in that way – hence the distinction between motivating and normative reasons. Justifiable reasons are also intended to be related to rationality – I argue for this in Chapters 4 and 9 – but justifiable reasons do not meet the requirements that Williams places on internal reasons.

An internalist account of normative reasons must both explain an agent's behaviour and be concerned with the agent's rationality. Assume that the mechanic in his friend's garage mixes

the stuff in the gin bottle with tonic and drinks it. Williams is right that a description of this agent's circumstances would explain his action, but, precisely because he had a false belief about drinking the stuff in the bottle, it would not show that he had good reason to do what he did. However, the agent's epistemic circumstances – the reasons he had this false belief – are relevant here. If the agent were acting rationally, he would have realised that the stuff might not be gin and checked his belief. This is not the case when agents have false beliefs that it is in practice impossible for them to correct. Assume that the blind man drinks the stuff in the gin bottle delivered to his house. He is not factually or rationally correctly informed. But it would be bizarre to claim that he was acting irrationally. On the contrary, it would have been irrational for him to phone a chemist to check the contents of the bottle.

Similarly, Smith argues that an agent's beliefs and reasoning must be corrected to determine what the agent has reason to do, because '[an] agent who has defective beliefs or who deliberates badly is indeed the sort of agent we tend to think of as being practically irrational' (1995a, p. 113). However, the words 'defective' and 'badly' are ambiguous. On one reading 'defective beliefs' and 'deliberating badly' are a sign of practical irrationality. Smith uses 'rationality' in this strong everything-must-be-gotten-right sense of the word (1996). On another reading of these phrases, an agent can be rational in spite of having 'defective beliefs' and 'deliberating badly'. If the mechanic decided to drink the stuff in the gin bottle, he would be practically irrational. His beliefs might be described as 'defective' and his deliberation 'bad', because he failed to appropriately investigate the facts. The blind man, however, came to have false beliefs in spite of appropriately investigating the world; his defective beliefs were not a sign of practical irrationality, and there is a sense in which he did not deliberate badly at all. Smith's argument does not show that taking account of epistemic circumstances when determining what an agent has reason to do fails to appropriately connect reasons to rationality in this second sense of the term.

My account of justifiable reasons is more appropriately associated with the rationality of the agent than Williams' and Smith's accounts of internal reasons, as it limits what an agent has reason to do to the set of actions that the agent can reasonably be expected to know she has reason to do. This was, in part, the very point behind Williams' development of internalism: 'internalism in general claims to show that it is a distinctive claim about *A* and *his* not φ-ing, that *A* has a reason to φ', (1995a, p. 44). It would have been rational for the mechanic to check his assumptions and reasoning. It would not have been rational for the blind man to do likewise. A description of the mechanic's situation, reasoning and motivational set would explain, but not justify, his drinking the petrol. A description of the blind man's situation, reasoning and motivational set would both explain and justify his drinking the petrol. An account of what an agent has reason to do that is based on everything that it is logically possible for any agent to take into account is focused on the rationality of the *action* rather than the rationality of the agent. In this latter case, what is being talked about is not that *A* has reason to φ, but rather that someone in *A*'s situation with perfect knowledge would have reason to φ.

7.2.3 Correcting false beliefs and faulty reasoning and learning relevant true beliefs corresponds to agents' own requirements on beliefs and reasoning

Williams argues that an agent's beliefs about what he is currently motivated to do must be corrected by adding all relevant true beliefs and correcting the agent's reasoning, because 'any rational deliberative agent has in his [subjective motivational set] a general interest in being factually and rationally correctly informed' (1995a, p. 37).[87] Williams' thought seems to be that the contents of an agent's motivational set that are held 'in general' will operate together with more particular focussed motivations to form the basis for those processed motivational states that are the basis for what agents have normative reason to do. Williams' claim that agents usually want to be factually and rationally correctly informed seems right, but this does not reveal a flaw in my account of justifiable reasons.

An agent who wants a gin and tonic will be motivated to ensure that his actions will lead to a mouthful of gin and tonic rather than a mouthful of petrol and tonic. Knowing what the stuff in the bottle is will increase his chances of drinking what he wants to drink. So, the agent does have an interest in being factually and rationally correctly informed. This does not mean that the agent only has normative reason to do something if the decision to act in that way is the result of correctly deliberating on all the relevant information. When agents have interests that compete with their desires to be factually and rationally correctly informed, they may be willing to shelve these desires when they are trying to determine how they ought to act.

When agents have competing interests, they also have competing reasons for acting. If a building has collapsed in an earthquake trapping people inside, the rescue workers will want to get people out as safely as possible. This will give them an interest in having all the correct information and reasoning skills that would enable them to do so. But, they will also want to get them out quickly. So, the rescue workers have a competing interest: they must decide which method to use to get people out without knowing whether their decision is based on a false belief, and without knowing which method is genuinely safest. Thus, the rescuers may have good reason to let their desire to save people quickly surmount their competing wish to ensure that they are rationally and factually correctly informed. Similarly, agents can have an interest in φ-ing when they are not factually and rationally correctly informed when it is inappropriate for them to act in ways that would lead them to be factually and rationally correctly informed.

The claim that agents have 'a general interest in being factually and rationally correctly informed' seems plausible. However, agents' competing interests sometimes take priority over this interest in being factually and rationally correctly informed, and it is appropriate that they do so. This argument of Williams' does not show that agents' interests in being

---

[87] Given the use that Williams makes of this 'general interest', I assume that he means an interest that applies generally across all actions, rather than an interest that applies sometimes, but not always.

factually and rationally correctly informed conflicts with the normativity of justifiable reasons.

7.2.4 Correcting false beliefs and faulty reasoning and learning relevant true beliefs ensures that informed agents will agree about what there is reason for a particular agent to do

According to Williams, it needs to be the case that '[what] we can correctly ascribe to [an agent] in a third-person internal reason statement is also what he can ascribe to himself as a result of deliberation' (1981a, p. 103; 1995a, p. 36). On my account of justifiable reasons, what an agent would have reason to do according to an observer may differ from what the agent has justifiable reason to do. Nevertheless, I hold that this difference does not introduce a worrying form of relativism into my account of justifiable reasons.

Assume that when an observer is deciding whether A has reason to φ, the observer only corrects A's beliefs in ways that it is possible and appropriate for A to correct them. In this case, a third-person account of what A has normative reason to do will match the conclusions that A would reach if she were to establish what she has justifiable reason to do. So, even though my account of justifiable reasons is based on a belief that practical and epistemic circumstances are relevant to normative reasons, what the agent judges she has justifiable reason to do will agree with a third-person account of what the agent has reason to do, as long as that third-person account applies an appropriate level of idealisation.

Limiting the extent to which A's beliefs are corrected does not introduce relativism of a kind that would have concerned Williams. For an observer to determine what agents have reason to do on Williams' account, the observer must understand the way in which the agent's motivational set functions. The move to adding that the observer must understand what it is appropriate for the agent to hold she has reason to do seems similar in kind. At times we will want to write about what an agent would have reason to do if she were in another set of epistemic circumstances. But, similarly, we sometimes want to talk about what an agent would have reason to do if other aspects of her circumstances were changed. And, this is not taken as a sign that facts about agents' circumstances are not relevant to what agents have reason to do. Someone might listen to an agent talk about her deliberations on some issue, know what pieces of information she needs to add into her reasoning, and know what she has reason to do before she has enough information to develop an intention to act in that way herself. The listener may even engage in reasoning that the agent should engage in herself, but in which she never engages. In other words, the listener may know better than the agent how she has reason to act. On my account an observer might even realise that there are good reasons for an agent to form an intention to φ when that agent could not in practice know she has reason to φ. As I argued earlier, the most extreme version of this third-person situation is the god-like view, where an omniscient observer knows all the relevant information for deciding what is best done in a particular situation. But note that, in the last two sentences, what the observer knows is not what *that agent* has reason to do, but what an agent with

greater knowledge would have reason to do. Such an approach is not consistent with Williams' project.

Smith also argues that our concept of normative reason is such that agents should agree on what there is reason to do in some set of circumstances. My account of justifiable reasons leaves me vulnerable to the charge of relativism that Smith makes when discussing Williams' work. Smith argues that reasons are relative on Williams' account, because '[Williams'] conception of reasons, like Hume's own, is predicated on skepticism about the scope for reasoned change in our desires' (1995a, p. 118).[88] The arguments in this thesis are based on scepticism about human agents' abilities to access the world and reason about those states of affairs of which they are aware. If an agent's justifiable reasons for action are limited by what it is appropriate for the agent to learn about the world and her motivational set, then there will be a sense in which an agent's justifiable reasons for action will be relative to her.

Although Smith claims that normative reasons are not relative, he acknowledges that normative reasons sometimes appear to be relative. However, he argues that the circumstances in which reasons appear relative all involve forms of relativism that are not worrying (Smith, 1995a, pp. 117-125). These forms of relativism arise because 'there is relativity built in to the *considerations* that we use to rationally justify our choices … not [because] … *our concept of a reason* is itself relative to the individual….' (Smith, 1994, p. 168; 1995a, pp. 120-121, italics in original). Two types of consideration lead to this non-worrisome form of relativity. The first consideration is based on the distinction between 'agent-relative' and 'agent-neutral' reasons (Smith, 1995a, p. 121). Suppose two children are drowning in a pond. Angela has more reason to save the one on the right, because he is her child. Bettina, who knows neither child, has no reason to save one rather than the other. This relativity is not worrying because if Bettina were in Angela's situation, she would have the same reason to save the child on the right. The second consideration Smith describes arises from agents having different preferences. If Angela likes gin, and Bettina loathes it, Angela may have reason to drink a gin and tonic when Bettina does not. This type of relativity is not worrisome, as it occurs when we take account of an aspect of Angela's (internal) circumstances that is relevant to what Angela has reason to do (Smith, 1995a, p. 122). Bettina could agree with Angela about whether Angela has reason to drink gin given her preferences (Smith, 1995a, pp. 122-123).

My account of justifiable reasons is compatible with this aspect of Smith's conceptual analysis of reasons. The relativity of reasons that results from my set of requirements also arises from taking account of considerations that justify our choices, rather than because the concept of a reason is made relative to individuals. According to Smith's account, agents have different reasons for acting if their circumstances are different, if, for example, Angela's child is among the two that appear to be drowning and Bettina's is not. In this case, the agents' reasons are not relative, because Angela and Bettina might agree that agents have reason to save their own child first. It is a short move from this view to the belief that agents can have

---

[88] This claim is based on Christine Korsgaard's arguments (Korsgaard, 1986).

different reasons for acting because they are in different practical and epistemic circumstances. On my account of justifiable reasons, agents' practical and epistemic circumstances are taken to be relevant to what those agents have reason to do, but it does not follow that justifiable reasons are relative to individual agents. What is relative to agents is their ability to obtain or process certain kinds of information. Even if practical and epistemic circumstances are taken to be relevant to whether A has reason to φ, judgements about what there is reason for a particular agent to do could converge.

The difference between my position and that of Williams and Smith arises because we use the word 'reason' in different ways: Williams and Smith are arguing in support of a form of objective normative reasons rather than justifiable reasons. The degree of idealisation they deem necessary to determine what an agent has reason to do puts at least some reasons into the realm of things that gods can know about rather than things that humans can always act on. However, I argue that we can gainfully distinguish between these two uses of 'reason', and that accepting an account of objective normative reasons does not require rejecting the concept of justifiable reasons. When we are talking about what a real human agent has practical normative reason to do, what the agent has reason to do is limited by the agent's practical and epistemic circumstances. When we are talking about what the agent would have normative reason to do if he was factually and rationally correctly informed, we are, in effect, trying to determine which action is ideal in some more general way. These two uses of the word 'reason' fit in well with the distinction, discussed above, between what A has reason to do and what there is reason to do in some set of circumstances.

The arguments that Williams and Smith give to show that the accounts of internal reasons that they develop are accounts of normative reasons emphasise the importance of basing any determination of what an agent has reason to do on relevant facts and correct reasoning. However, these arguments do not show that justifiable reasons are not normative, nor do they reveal any other flaws in my account of justifiable reasons. Interestingly, two arguments that Williams gives that appear to be directed against accounts of reasons such as mine can be used in support of my position. A deeper consideration of Williams' motivation in developing his internalist theory of reasons suggests that Williams would have been sympathetic to my position.

7.3 Bernard Williams and justifiable reasons

My initial thought that Bernard Williams' internalism might offer a theory of reasons for action that is responsive to the nature of real human agents was incorrect, and some of Williams' arguments suggest that he would not have accepted my account of justifiable reasons. Williams also knew that his internalist account of reasons entailed that agents might sometimes not be able to know what they have reason to do. He writes:

> [T]he internalist can give an account of ['If A deliberated correctly, he would be motivated to φ'] in which it may be replaced by ['A has a reason to φ'] …. But this does not imply … that the agent should be able to conduct the relevant deliberation in fact. Perhaps some unconscious obstacle, for instance, would

have to be removed before he could arrive at the motivation to φ (Williams, 1995b, p. 188).

Williams claims that 'unconscious obstacles' can limit agents' abilities to learn that they could be motivated to act in a certain way. Someone who has a phobia, for example, might not be able to carry out the reasoning necessary for her to learn that, with psychological work, she could be motivated to act in ways that conflict with her phobia. This means that she might never be able to learn that she has reason to act in ways that conflict with her phobia. Other aspects of agents' circumstances will also affect what they can learn about what they have reason to do. This acknowledgment that his theory entails that agents might not be able to learn what they have reason to do and his insistence that reasons are identified with reference to actual states of the world are grounds for doubting that Williams would accept my account of justifiable reasons. However, some of the arguments Williams gives in support of internalism suggest that Williams was focussed on developing a theory that explained real agents' reasons for acting, which in turn suggests that he would have been sympathetic to my account of reasons.

Bernard Williams' arguments in §7.2 are all ostensibly opposed to accounts of reasons that are not based on facts about agents' circumstances. My discussion of these objections shows that they do not impact on justifiable reasons, but it also shows that my account of justifiable reasons is better at achieving some of Williams' goals than his own theory. First, Williams claims that beliefs and reasoning need to be corrected to determine what an agent has reason to do because such corrections correctly link reasons and rationality. However, limiting any correction of beliefs and reasoning to corrections that it is appropriate for agents to carry out produces a more intuitively plausible relationship between reasons and rationality than the relationship that arises from correcting beliefs and reasoning in ways that are impossible for the agents themselves to carry out. As I claimed earlier, the blind man with no sense of smell would not be thought rational if he called in a chemist to test the contents of the gin bottle, and might be thought irrational if he did not drink the petrol in the gin bottle. Second, Williams mentions that agents themselves want to be factually and rationally correctly informed, and it seems reasonable to accept this claim. However, it also seems reasonable to assume that agents generally have in their motivational sets a desire to make the best decisions possible in their circumstances. Making the best decisions possible in their circumstances will not involve taking things to be reasons when they are the result of factually and rationally ideal information. It will involve determining what the agent has justifiable reason to carry out. These two arguments of Williams' seem to support my theory of justifiable reasons more than his own development of internal reasons.

Williams' emphasis on the importance of reasons being related to actual agents rather than highly idealised abstract agents also seems to support my theory of justifiable reasons. Williams rejects theories of reasons that he takes to claim that 'A has a reason to φ' means 'if x were a correct deliberator, x would be motivated in these circumstances to φ'(1995b, p. 189). He writes that such a statement is not 'distinctively about A at all' (Williams, 1995b, p. 189). In his articles on internal reasons, Williams' focus is always on the agent, and what that agent has normative reason to do, rather than on what there would be reason for some

idealised agent to do. Williams writes: 'internalism in general claims to show that it is a distinctive claim about *A* and *his* not *φ*-ing, that *A* has a reason to *φ*' (1981a, p. 101; 1995a, p. 44 , quote, italics in original). And in his reply to comments made by McDowell, Williams writes:

> On the externalist account, so construed, statements of the type [*A* has a reason to *φ*] do not relate actions to persons, but types of actions to types of circumstances, and they are most revealingly expressed in the form 'in circumstances *X*, there is reason to *φ*.' (Williams, 1995b, p. 190, see also 191-192)[89]

According to Williams' version of internalism, the reasons that individual human agents have for acting will arise partly from the nature of those particular agents. A particular human agent may have reason to act in ways that an abstract correct deliberator would never have reason to act in, simply because the abstract agent's ability to deliberate is so highly idealised.

Williams' concern that some theories of practical reason are based on ideals rather than the nature of human agents was part of his broader concern that utilitarian and Kantian moral theories abstract too much from the nature of particular agents (1981b). Williams claims that the ideals used by utilitarians and Kantians in their theories take the focus off the reasons that individuals have to pursue projects that are of value to them, projects that are based in individuals' desires and concerns and which give people's lives purpose (1981b, pp. 5, 12). According to Williams, Kantians err by basing their claims about what people have reason to do on abstract agents. He argues that on a Kantian understanding of reasons, people can be required to give up projects that are key to them having an interest in their own existence so that they can facilitate 'the impartial good ordering of the world of moral agents' (Williams, 1981b, p. 14). Williams claims that utilitarians err by thinking in terms of amalgamations of agents, and he argues that this explains why utilitarians require people to give up their projects if those projects do not appropriately relate to the overall good (1981b, p. 14). In more directed comments, he criticises the work of John Rawls and Thomas Nagel for requiring that agents take 'an external view of … [their lives], as something like a given rectangle that has to be optimally filled in' (Williams, 1981b, p. 12). While Williams' main focus here is moral motivation, his theory of reasons connects moral motivation to theories of reasons, and his emphasis on the importance of reasons being appropriately related to particular agents seems likely to be linked to these arguments about the importance of making reasons connect to agents' projects.

Williams' arguments that moral theories should take account of the importance of real agents' projects is paralleled by his arguments that the deliberation expected of agents be limited to deliberation that agents could practically undertake. Williams argues that the forms of reasoning expected of agents include such forms of deliberation as the 'effective direction of … attention', 'thinking of alternatives', '[perceiving] unexpected similarities', and using imagination to resolve conflicts (1981a, pp. 102-103, 105, 108; 1995a, pp. 36, 38, 42). These are all practical approaches to reasoning that can be, and are, carried out by limited human

---

[89] Williams makes the same point in his replies to Scanlon (2001, p. 96).

agents. I take it that Williams would agree that the more extreme the idealisation of agents' reasoning abilities used to ascertain an agent's reasons for acting, the less the resulting reasons are about the particular agent.

The claim that Williams' understanding of deliberation is limited in this way is reinforced by McDowell. According to McDowell, Williams holds that an agent's motivational set limits the forms of deliberation that can be used to establish what that agent has reason to do (McDowell, 2006, p. 172). I take this to mean that the contents of an agent's motivational set include things that affect the forms of reasoning that an agent can be motivated to undertake. Such a limitation will restrict agents' normative reasons in ways that cause them to resemble justifiable reasons. An agent's motivational set will constrain the ways in which the agent's existing set of beliefs can be corrected to determine what the agent has reason to do. If, for example, someone finds all reasoning that uses symbols so difficult and painful that it is not appropriate to expect him to engage in such reasoning, any ignorance or false beliefs he has that can only be corrected through his careful study of symbolic logic will not be corrected when determining how he has reason to act. This suggests that Williams' criteria for deliberation push his position close to my account of reasons.

Williams intends to develop an account of reasons that relates to specific agents. This suggests that my focus on limiting what an agent has reason to do to what it is appropriate for the agent to hold he has reason to do is consistent with Williams' general project. If McDowell is right that Williams holds that an agent's motivational set limits the forms of deliberation that can be used to establish what that agent has reason to do, then, given that I make the same claim about justifiable reasons, this brings Williams theory more into agreement with my account of justifiable reasons. I suspect that Williams' assumptions about what can serve as reason-explanans in the *STATE* or *END* senses of the word conflict with my approach because Williams is focussed on the issue of value rather than on the *STATE* and *END* aspect of reason relations.

7.4 Conclusion

Arguments about whether reasons are internal or external, objective or subjective, or whether there can be justifiable reasons, are all arguments about the nature of reason relations. However, different aspects of reason relations are at issue in these arguments. Internal-external reasons arguments focus on the nature of the values that underpin reason relations. In contrast, the arguments about the objective-subjective distinction that are discussed in this thesis focus on the way in which beliefs, propositions or states of affairs come to serve as reasons for action. This makes it unproblematic for the accounts of internal reasons given by Williams and Smith to be considered accounts of objective normative reasons. Showing that an account of internal reasons can also be an account of objective normative reasons is not sufficient for showing that justifiable reasons are not internal reasons. However, if there are external reasons, nothing about justifiable reasons suggests that they must be internal reasons rather than external reasons. So, I also showed in this chapter that although what an agent has

justifiable reason to do will be affected by the agent's motivational set, my account of justifiable reasons does not assume the truth of internalism or externalism about reasons. I showed, therefore, that the debate about whether internalism or externalism is true addresses a different issue from the debate about whether there are justifiable reasons.

Several of Bernard Williams' and Michael Smith's arguments in support of internalism appear to be arguments against my account of justifiable reasons. I argued that these arguments do not entail that normative reasons can only arise from taking a gods'-eye perspective on the world. Williams is correct to argue that his sound deliberative route achieves three desiderata: it makes normative reasons normative; it ensures such reasons are appropriately connected to an agent's rationality; and it ensures that normative reasons are, in a sense, non-relative. Williams is not, however, correct to claim that an agent who wants a gin and tonic cannot have reason to drink a petrol and tonic. A deliberative route that is limited by what it is practically possible for an agent to learn can also provide an account of internal reasons that achieves the three desiderata.

Finally, I argued that there are grounds for thinking that although Williams gives an internalist account of objective normative reasons, there is much about his arguments about internalism and ethics that is consistent with my account of justifiable reasons. In spite of the audacity of the claim, I suspect that he would be sympathetic towards my account of justifiable reasons. Williams' position is clearly an unusual one for objective normative reasons theorists. In the next chapter, I look at accounts of objective normative reasons in more depth.

8. Objective Normative Reasons

8.0 Introduction

Accounts of objective normative reasons serve a different function from my account of justifiable reasons. The accounts of objective normative reasons discussed in this thesis are accounts of reasons that produce the right results or that realise the relevant values, where 'values' is read in a broad, inclusive way. Developing an account of the reasons that mean that an action will produce the right results or realise the relevant values is important. Accounts of reasons that produce the right results or realise relevant values are a step towards developing an account of value, and they allow us to make sense of statements such as: 'If only he could have known what he had reason to do'. This chapter provides arguments that support claims made about objective normative reasons made earlier in this thesis. It does not provide a critical analysis of accounts of objective normative reasons; it considers the function of objective normative reasons and the forms of idealisation used to specify when a reason counts as an objective normative reason.

The first section in this chapter has a different focus from the rest of the chapter, although it also serves as an initial illustration of the function of objective normative reasons. Throughout the thesis I claim that accounts of objective normative reasons are not concerned with the accessibility of reasons to real human agents, but provide little evidence to support this claim. So, in §8.1, I use two positions on objective normative reasons that I have not yet examined in this thesis, those of Torbjörn Tännsjö and Joshua Gert, to explain some of the grounds objective normative reasons theorists give for ignoring the accessibility of such reasons (Gert, 2008; Tännsjö, 2010).

Earlier in the thesis I claimed that there are no problems with accepting both an account of objective normative reasons and an account of justifiable reasons because the two forms of reasons serve different functions.[90] In §8.2, I explain some of the functions of objective normative reasons, functions that justifiable reasons cannot serve, and discuss some of the ways in which such reasons are derived from idealising about agents, the world or reasoning. Objective normative reasons complement rather than compete with justifiable reasons. To show this, I use Michael Smith's work in *The Moral Problem* as an extended example of the function of objective normative reasons and the ways in which they differ from justifiable reasons. Other theorists develop different accounts of objective normative reasons that play different roles within philosophers' research programmes. I explain the positions of several philosophers with very different approaches to Smith's to show that although they use objective normative reasons in the development of very different theories, objective normative reasons serve similar functions within each theory. The conceptions of objective

---

[90]This was initially discussed in §5.3.

normative reasons that these philosophers develop vary due to the different theories in which these reasons play key roles.

In §8.3, I explain two ways in which accounts of objective normative reasons can help real human agents determine what they have justifiable reason to do. When an agent has a justifiable belief that she has an objective normative reason to act in some way, she has a justifiable reason to act in that way. This means that accounts of objective normative reasons could sometimes serve as guides to when an agent has justifiable reason to act in some way. Second, my account of justifiable reasons makes no claims about what has value or what might count as an end that is worth pursuing. If an account of objective normative reasons successfully explains what serves the role of *VAL* and *SIG* within reason relations, this would help develop people's understanding of the kinds of things that they should take themselves to have justifiable reason to do.[91]

The frequent mention of 'right results', 'value', and 'right action' in this chapter is easily read as implying that the reasons at issue are overall normative reasons rather than pro tanto normative reasons, but this is not the case. Unless I clearly state that I am referring to overall right results or the action that is most valuable, take 'right results' to refer to results that are in some way and to some degree right, and take 'value' to refer to what may be one value among many. Similarly, unless stated otherwise, 'right action' does not refer to the action that is right all-things-considered. Take 'right action' to refer to an action that has something to be said in favour of its being carried out, that is, an action that is in some way or to some degree right.

8.1 The irrelevance of the accessibility of objective normative reasons

The accessibility of objective normative reasons to real human agents is usually thought irrelevant by objective normative reasons theorists. Some of those who give accounts of objective normative reasons explicitly state that such reasons may be inaccessible (Williams, 1995b, p. 188). But, although most of the accounts of objective normative reasons that I have read ignore their accessibility to real human agents, not all accounts of objective normative reasons ignore their accessibility. Jonathan Dancy is a notable exception. Dancy introduces the idea of an epistemic filter that limits what an agent can have objective normative reason to do to things that the agent could determine she had normative reason to do.[92] He writes that 'the grounds for our reasons, like the reasons themselves, must lie within our capacities for recognition, if they are to be capable of being practically relevant for us' (Dancy, 2000, p. 59). And, like me, Dancy aims to develop a theory of practical reason that is practically relevant

---

[91] *VAL* and *SIG* and their role within the reason relation are discussed in §3.2.

[92] In the Normative Reasons Venn Diagram in the next section, Dancy's objective normative reasons occupy the intersection between the sets labelled 'objective normative reasons' and 'justifiable reasons'. This amounts to a claim that all objective normative reasons are accessible, not a claim that all objective normative reasons are justifiable reasons. See §5.3 for more information.

to agents. Dancy's approach is, however, atypical. To show some of the ways in which objective normative reasons are thought inaccessible to agents, I briefly describe two accounts of reasons that have not yet been mentioned, those of Torbjörn Tännsjö and Joshua Gert (Gert, 2008; Tännsjö, 2010). I consider Tännsjö's account because he takes an extreme position, arguing that objective normative reasons may never be accessible to agents. Gert's position is more standard, but still entails that few agents ever act for objective normative reasons.

Tännsjö claims that we need two concepts of reasons, which he calls 'Humean reasons' and 'moral reasons' (2010, pp. 11, 27-29, 75-90, 152-154). Neither his 'Humean reasons' nor his 'moral reasons' resemble justifiable reasons. Tännsjö's 'Humean reasons' are subjective reasons based on agents' beliefs and desires rather than on facts and principles. His 'moral reasons' are a form of objective normative reason that includes moral, aesthetic, and prudential normative reasons.

> [Moral] reasons are abstract true propositions (facts) capable of explaining a normative fact.… Moral reasons form the premises (the explanans) of an inference *to a normative fact* (the explanandum).… [A moral reason is an explanation if it includes] essentially among the premises a *moral principle.* Moral principles play, then, the same role in moral explanations that laws of nature play in scientific explanations of a standard 'Hempelian' model. (Tännsjö, 2010, pp. 27-28)

For Tännsjö, the normativity of moral reasons comes from the moral principles that he claims are pre-requisites for all moral reasons (2010, p. 29). Tännsjö argues that his 'moral reasons' are often, or perhaps always, undetectable by real human agents:

> To me it seems perfectly in accordance with moral phenomenology as it actually is to acknowledge that, in many situations, even if we have done our best to find out what to do, we may have failed. I would be surprised to find that our moral obligations were ever quite accessible to us. (Tännsjö, 2010, p. 33)

Tännsjö bases his claim that normative reasons are inaccessible to agents on two arguments. I find the first unconvincing, but the second plausible.

First, Tännsjö argues that normative reasons never explain agents' actions because agents' actions must be based solely on beliefs rather than on states of the world (2010, pp. 37-38). So, using the terms in Tännsjö's way, motivating reasons must always be Humean reasons and are never moral reasons. I am not convinced by this claim. As explained in earlier chapters, I accept Dancy's argument in *Practical Reality* that it is usually appropriate to take reality to be the source of practical reasons; my reason for moving my fingers as I do right now is the structure and function of the computer keyboard, not my beliefs about the computer keyboard (2000). This issue is discussed further in §9.5.

Tännsjö's second argument for the inaccessibility of objective normative reasons arises from his approach to ethics (2010, pp. 32, 33-34). Tännsjö argues that moral reasons need not be practically relevant:

> [We] have no guarantee that moral truth must (always) be applicable in practical life. The assumption that it must is based on an unwarrantedly simplistic picture of our moral predicament…. [A] moral principle need only guide choices in the sense that an omniscient deliberator should be able to apply it. (Tännsjö, 2010, pp. 33-34)

Tännsjö claims that agents ought to do whatever will bring about what is of value, irrespective of whether they can learn what has value or learn how to realise that value (2010, pp. 32, 33-34, 90-91). He suggests that our knowledge of our circumstances and of the consequences of our actions falls so far below the ideal that we can never know what we have reason to do: 'We cannot comprehend all the consequences of our actions, let alone can we survey what would have happened, had we acted differently' (Tännsjö, 2010, p. 90). This approach to ethics is, as Tännsjö claims, consistent with some forms of consequentialism and some ideal observer theories (2010, pp. 32, 34, 90). And, as he points out, this way of thinking about good reasons for actions allows us to make sense of sentences such as, 'He has reason to bet on the winning horse even though he cannot know which horse will win.' (Tännsjö, 2010, p. 33). Thus, although his objective normative reasons are inaccessible to real human agents, Tännsjö takes himself to provide an account of normative reasons that serves an important function, a function that objective normative reasons could not serve if such reasons were limited in a way that made them always accessible to real human agents. This function is discussed further in the next section.

Joshua Gert refers to 'objective practical reasons', rather than 'objective normative reasons', but the terms are synonymous (2008). Gert argues that objective practical reasons exist independently of any agent's ability to determine that they exist (2008, pp. 316-317). As an example of objective practical reasons, he cites the fact that an 'action will increase (or decrease) the risk of someone avoiding (or suffering) a harm such as pain' (Gert, 2008, p. 317). He writes:

> Suppose that a certain act increases one's chances of suffering a painful illness. [This] … typically counts as a reason against the action – whether or not one is aware of it. And this means that we can reasonably say that people in 14th century France were typically unaware of all the reasons they had not to pile up dead bodies near food markets.… But time cannot change the fact that we have reasons – not always decisive, of course – to avoid death and pain, and to seek knowledge and pleasure. (Gert, 2008, p. 322)[93]

Gert's approach to reasons differs from that of Tännsjö, but his claims about 'objective practical reasons' are similar to Tännsjö's claims about his 'moral reasons'. Like Tännsjö, Gert claims that irrespective of what we can know about the world, the way the world is, and the way that we are, gives us objective normative reasons to act in certain ways.

On Gert's account of objective practical reasons, real human agents will often not be in a position to know that they have reason to act in a certain way. Just as we know that the

---

[93] Note, only for the sake of interest, that Gert uses the word 'reason' in two ways in this passage. When he writes 'that a certain act increases one's chances of suffering a painful illness … typically counts as a reason against the action', the word 'reason' is used to refer to what I call '*END*'. In the second part of the quote he uses 'reason' to refer to the reason relation as a whole.

world-view of people in fourteenth century France was extraordinarily limited, we know that seven centuries from now our understanding of our circumstances and the consequences of our actions will appear extraordinarily limited. But the inaccessibility of objective normative reasons goes beyond an inability to know what we have objective normative reason to do. Our limited understanding of the world and ourselves means that agents are commonly not even in a position to act on Gert's objective practical reasons.

Although the features of Tännsjö's, Gert's, and other theorists' accounts of normative reasons ignore their accessibility to agents, these features also give objective normative reasons a function that supports the need for an account of objective normative reasons without undermining my claims about the importance of justifiable reasons. As Tännsjö and Gert claim, the following are all plausible claims about reasons:

> 'He has reason to bet on the winning horse even though he cannot know which horse will win.'

> '[People] in 14th century France were … unaware of all the reasons they had not to pile up dead bodies near food markets.'

> '[We] have reasons – not always decisive, of course – to avoid death and pain, and to seek knowledge and pleasure.'

> 'If only I could have known I had reason to φ.'

These claims are all about reasons that the relevant agents could not consistently use to guide their actions; in this way they are all impractical reasons. So, what are these reason claims about?

8.2 What function do objective normative reasons serve?

Normative reasons are reasons that agents *really* have some reason to act on; normative reasons cannot be something that has been *mistaken* for a reason. But the claim that an agent really has reason to act in some way can be read in two different ways, and the different functions of justifiable reasons and objective normative reasons are apparent in these different readings of the phrase. On the justifiable reason way of reading the phrase, it is a reference to the ways in which an agent has some reason to act. The normativity, and idealisation, of such reasons comes from the role of justification in the reason relation. On the objective normative reason way of reading the phrase, it is, roughly, a reference to the potential value of an action. The normativity, and idealisation, of such reasons comes from the role of results or value in the reason relation.[94] I write 'roughly, a reference to the potential value of an action' because I intend this claim about the relationship between objective normative reasons and values to be read as broadly as possible. In particular, as in §3.2, I intend it to have no implications for the truth or otherwise of buck-passing accounts of value. The word 'value' must also be read broadly. An action can have value because of the kind of action it is rather than its outcome. On this way of thinking about value, there is

---

[94] In other words, the normativity and idealisation of such reasons is derived from the characterisation of '*END*', '*VAL*' and '*SIG*'. See Chapter 3, and below, for more details.

nothing odd about saying that an agent's acting from goodwill has value, irrespective of the consequences of the action. The word 'value' must also be read in a way that does not presume that all values are moral values, or that for an action to have value it must make a major mark on the world. There can be value in little things, such as stepping on a spider or not stepping on a spider, holding a violin just so, or getting out of bed on the right side rather than the left. Although the values that play a role in theories of objective normative reasons need not be moral values, theories about moral reasons often lie behind the development of those theories. So, I briefly explain the relationship between objective normative reasons and moral reasons before continuing on to explain the function of objective normative reasons given by Michael Smith in *The Moral Problem* (1994). Finally, I show that accounts of objective normative reasons that conflict with Smith's serve the same general function that I argue is served by objective normative reasons.

Objective normative reasons are accounts of reasons that are correctly related to potential outcomes, whether those potential outcomes consist of the correct performance of a particular kind of action or consist of increasing value in some other way. What an agent has an objective normative reason to do is determined by taking an idealised standpoint when considering the state of the world, agents, or forms of reasoning. As is already apparent, the form of idealisation involved, and what gets idealised, varies from theorist to theorist. I argue that:

> *A* has an objective normative reason to φ iff for *A* in *C*, *STATE, END, VAL,*
> and *SIG*.

As described in §3.2, '*STATE*' is a state of affairs, '*END*' states that given that state of affairs, a particular action will have a particular result, '*VAL*' states that the result has positive value, and '*SIG*' states that the value is not significantly outweighed by negative values that may arise from the agent's φ-ing. If objective normative reasons theorists were to accept my claim that reason relations include the elements *STATE, END, VAL*, and *SIG*, then different objective normative reasons theorists could be said to offer different accounts of the form of idealisation required to identify the correct *STATE, END, VAL*, and *SIG*.[95] What an agent has justifiable reason to do is also determined by taking an idealised standpoint. Justifiable reasons are derived from idealisations about what particular agents in particular circumstances would be justified in taking themselves to have reason to do. This explains why it is correct for real human agents to take themselves to have reason to act as they have justifiable reason to act.

Although debates about what it means for reasons to be appropriately connected to value are important for all forms of reasons, including prudential reasons and aesthetic reasons,

---

[95] Mark Schroeder's rejection of what he calls the 'No Background Conditions' view of reasons means that he might reject my analysis of reason relations in terms of *STATE, END, VAL* and *SIG* (2008c, pp. 23-40). Schroeder might claim that including *VAL* as part of the reason relation is a step too far towards saying that an agent's desire for some end can be why the agent has an objective normative reason for acting. But, given Schroeder's use of 'desire' within his reason relation (see, 2008c, p. 59, for example), it is just as plausible that he would accept my analysis of the reason relation, but reject my claim, in §3.2, that it can be appropriate to use the term 'reason' to refer to *VAL* or *SIG*.

concerns about moral reasons and values often lie behind debates about objective normative reasons. The forms of idealisation used to develop accounts of objective normative reasons vary with the theories about ethics accepted by the philosopher. Although accounts of objective normative reasons are never intended to apply only to ethics and moral values, every account of objective normative reasons considered in this thesis was developed as part of a search for an account of practical reason that agrees with, and supports, particular accounts of ethics and value. The claim that these philosophers' ethical standpoints affect their accounts of reasons and vice versa is not a criticism. Theories in such tightly connected fields need to be developed in concert, with conclusions about one field providing starting points for the other, and with positions in both fields altered after reflection.

Practical reason and philosophy of action may seem distant from meta-ethics and normative ethics, but moral reasons are a species of normative reasons, so claims about the nature of normative reasons impact on the plausibility of claims about ethics. The link between the development of theories about normative reasons, values and ethics is unsurprising. Consider the diagram below.

Normative reasons Venn diagram
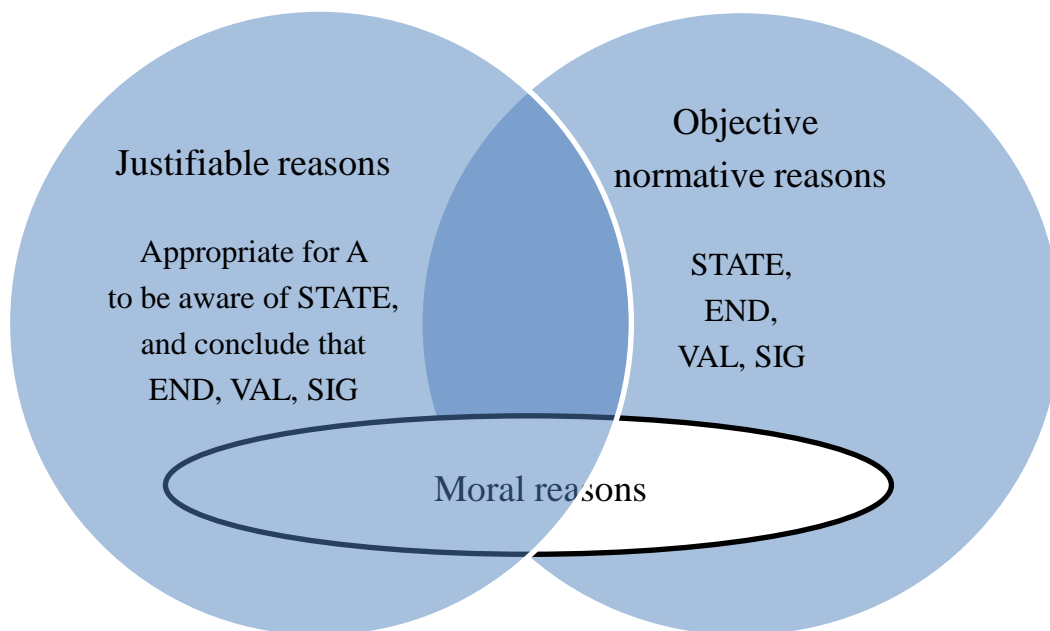
Values, in the broad sense of the term, help determine what agents have normative reason to do; this explains their role within normative reason relations. Moral reasons are reasons for action; they are a subset of normative reasons. The nature of the values operating within a reason relation seems likely to be what determines whether normative reasons are moral or non-moral reasons. And, the nature of the values that underpin moral reasons will affect the plausibility of any ethical theory. For example, consequentialists and deontologists offer different accounts of what agents have moral reason to do because of differences in their

accounts of moral reasons and their positions on the sources of value. Philosophers' awareness of this relationship between practical reason and ethics explains the ways in which their meta-ethical standpoints affect their accounts of normative reasons, and, in particular, their views of the relationship between reasons and values.[96]

My claims about the distinction between objective normative moral reasons and justifiable moral reasons can easily be illustrated using examples that involve wellbeing or other common consequentialist values, such as the case of the collapsed stranger, but Kantian examples can also be used. Consider Kant's position on lying. Kant argues that lying is always wrong because it violates the categorical imperative ([1785] 1993; [1797] 1993). Imagine that someone is threatening to cause great harm to your children. If you refuse to reveal where your children are, your partner will be killed in front of you. Lying will buy you more time, which may mean that help arrives. In the reason relation for this case:

STATE: Someone has asked you where your children are and threatened to kill your partner if you don't say where your children are.

END: If you lie, you may buy enough time for help to arrive.

VAL: Lying will violate the categorical imperative.

SIG: Kant states that potential bad (or good) consequences cannot outweigh the categorical imperative.

As described, the agent has a Kantian objective normative moral reason to tell the truth or say nothing. However, the agent might reach a justifiable, but mistaken, conclusion about *VAL* or *SIG*. Perhaps, after reasoning to the best of her ability, the agent concludes that the maxim 'lie to psychopaths when lying might prevent grievous harm' agrees with Kant's categorical imperative. If she lies, then she does so for a justifiable moral reason that is not a Kantian objective normative moral reason (Kant, [1785] 1993; Kant, [1797] 1993).

Michael Smith's development of an account of objective normative reasons illustrates the way in which an account of objective normative reasons can be used to develop and support accounts of normative ethics, meta-ethics and values. What follows is an exposition of Smith's position, not a critical evaluation; my claim is not that Smith provides an account of objective normative reasons that succeeds in forming a foundation for ethical theorising or for determining what has value, but that his account of objective normative reasons shows how an account of objective normative reasons can serve such a purpose. The exposition should be read as an account of pro tanto reasons rather than an account of overall reasons, as

---

[96] The 'Normative reasons Venn diagram' has one implication that I disagree with, although my disagreement can be treated as a disagreement about terminology. I find it natural to use the term 'moral reasons' to refer to justifiable moral reasons – the shaded section of the 'moral reasons' set above. I consider the members of the set of 'moral reasons' in the white section above, that is, 'moral reasons' that are not justifiable reasons, as indicators of positive value, not indicators of right action. Agents cannot use the reasons in this section of the Venn diagram to guide their actions. So, I hold that there is no sense in which it would be right for agents to act on them. This assumption does not conflict with consequentialist theories that claim that agents ought not to focus on the consequences of their actions. I take it that according to such theories agents can still act in response to moral reasons when they uphold individuals' rights, for example. This assumption is also consistent with the common assumption that we may have moral obligations that we are not motivated to act on.

Smith intended it be read in this way unless otherwise stated (1994; 1996, p. 167; 1997a, p. 92).

In *The Moral Problem*, Michael Smith sets out what he calls 'the central organizing problem in contemporary meta-ethics', a problem that is the focus of many arguments about cognitivism and non-cognitivism, realism and anti-realism, and about Humean accounts of reasons and motivation (1994, p. 11).[97] This problem arises from the plausibility of three apparently inconsistent statements:

1. Moral judgements of the form 'It is right that I *φ*' express a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.[98]

2. If someone judges that it is right that she *φ*s, then *ceteris paribus*, she is motivated to *φ*.

3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences. (Smith, 1994, p. 12)[99]

The conflict between these statements arises because the truth of any two of them seems to conflict with the truth of the third. If (1) and (2) are true, then moral judgements are beliefs and those beliefs motivate agents. But this seems to conflict with (3) which states that desires motivate. If (1) and (3) are true, then desires motivate agents, and moral judgements are beliefs, but this conflicts with (2) which claims that moral judgements motivate agents. If (2) and (3) are true, then desires motivate and moral judgements express desires, but this conflicts with (1) which states that judgements are beliefs. Rather than rejecting one of these statements, Smith argues that they can be reconciled, and his analysis of normative reasons is key to this reconciliation.

Smith claims that what it means for an action to have value and to be morally right can be explained using the correct account of normative reasons (1994, pp. 130-181, 182-184; 1997a, pp. 88, 107, 117). Smith equates valuing an action with holding that there is normative reason to perform that action. He argues that it is morally right for an agent to act in some way when two conditions obtain. First, the agent has a normative reason to act in that way. Second, the action is in the realm of the moral, where the 'realm of the moral' includes, for example, situations where respect for others and human flourishing are at issue. So, on Smith's account

---

[97] Debates between objective normative reasons theorists tend to focus on the relationship between value and the idealisation necessary to establish what counts as a reason for action, and many of these debates focus partly on the Humean account of reasons. In brief, on the Humean theory of reasons, normative reasons are either explicitly linked to agents' desires, or they are explained instrumentally, that is, explained in terms of agents' pre-existing goals. So, on a Humean theory of reasons, it is desires, or the satisfaction of pre-existing goals, that serve the role of value in *VAL* and *SIG* in the reason relation. This means that by setting aside debates about the Humean theory of reasons, I have set aside arguments that are important for the positions of many of those discussed in this thesis. These arguments have been set aside because they are tangential to my concerns in this thesis.

[98] Smith takes seriously the possibility that we are all mistaken about what we think we are doing when we make moral judgements, but (I think correctly) does not take this to show that there is no point in arguing as he does (1994, pp. 187-202).

[99] See also, (Smith, 1989, pp. 89-92).

of moral judgements, the normative reasons that agents have moral reason to act on are connected to objective facts about what agents have reason to do. Smith needs to develop an account of normative reasons that is related to objective facts in a way that can make agents' actions morally right. By doing so, he can show that there are grounds for holding that agents making moral judgements are not deluded about the existence of moral reasons or deluded about what they have moral reason to do.

Smith argues that an agent has a normative reason to act in some way when she would be at least somewhat motivated to act in that way if she were fully rational (1994, p. 181).[100] He takes the claim that 'what we have normative reason to do is what we would desire to do if we were fully rational' to be a platitude (Smith, 1994, p. 150).[101] For an agent to be 'fully rational', as Smith uses the term:

> (i) the agent must have no false beliefs
>
> (ii) the agent must have all relevant true beliefs
>
> (iii) the agent must deliberate correctly (Smith, 1994, p. 156)

Smith writes that he adopts these conditions from Bernard Williams' account of normative reasons and he refines them by developing an account of what correct deliberation requires that differs from that of Williams (Smith, 1994, pp. 156-161). The requirement that normative reasons are not based on 'false beliefs' and are based on 'all relevant true beliefs' is one step towards ensuring that normative reasons are based on objective facts, but does not on its own connect Smith's account of objective normative reasons to values. Smith's account of correct deliberation is also important. He argues that agents deliberate correctly when they submit their sets of desires to a process of systematic justification. This systematic justification leads agents to discard some desires and form new desires, until they come to have a fully rational set of desires. He claims that systematic justification would lead all agents to agree about what they desire in a particular set of circumstances (Smith, 1994, pp. 164-174). Indeed, on Smith's account of 'full rationality', if agents do not all agree about what there is normative reason to do, there is no normative reason to act in that way (1994, p. 198). So, according to Smith, if any agent has a normative reason to act in some way, all agents in the same circumstances have reason to act in that way. Any vagaries that stem from individual differences become unimportant. This means that moral reasons are not relative in any way that might undermine ethics (Smith, 1994, pp. 187-189, 193-202).

The way in which Michael Smith abstracts away from individual limitations is made clear in his reply to a criticism from Christine Swanton (Smith, 1996; Swanton, 1996). Swanton suggests that people might satisfy Smith's requirements (i) – (iii) and yet hold different views about what there is normative reason to do. Swanton argues that fully rational optimists might hold that they have normative moral reasons to have children, while fully rational people who

---

[100] Smith's account of normative reasons is introduced in §5.2.2 and §7.1.1.

[101] Smith modifies this position by distinguishing between the kinds of advice different fully rational advisers would give an agent about what the agent has reason to do. This modification is important for the plausibility of his theory, but is not relevant here. See my discussion of Exemplar and Adviser in §5.2.2.

are (non-clinically) depressed about the state of the world would regard it as immoral to have children (1996, p. 158). If fully rational people have different views about what there is normative reason to do, then, because of the way in which Smith connects normative reasons and the desirability of actions, the desirability of actions is relative to individuals. However, Smith rejects Swanton's claim on the grounds that:

> fully rational agents, as I have characterised them, are unable to be either optimistic or pessimistic, because fully rational agents have all the information that there is, where this includes, a fortiori, information about how events in fact turn out. Optimism and pessimism are thus simply not dispositions that fully rational agents can so much as possess, but are rather dispositions that only less than fully rational agents – agents who are at least informationally deprived – can possess. (Smith, 1996, p. 166)

So, Smith considers that to be fully rational, agents must know 'how events in fact turn out'. This helps ensure that all agents will agree about what there is objective normative reason to do, and, hence, agree about what is desirable, and what has value, but real human agents do not know how events will turn out. Real human agents cannot, therefore, be fully rational, and Smith is aware of this implication of his position.

Smith's analysis of normative reasons allows him to solve the moral problem by reconciling the three statements that produce the problem. To recap, Smith explains value and right action in terms of normative reasons, and normative reasons in terms of what a fully rational agent would believe and desire, where a fully rational agent reasons from the relevant facts and has a coherent set of beliefs and desires. Smith's reconciliation of the three statements that lead to the moral problem modifies the first two statements slightly (1994, pp. 184-185):

$1_{rational}$ Moral judgements express beliefs about what an agent would desire if she were fully rational.

$2_{rational}$ If someone judges that it is right that she φ-s, then, if she is rational, she will be motivated to φ.

3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences.

So, when an agent judges that an action would be morally right: she expresses a belief that she would want to act in that way if she were fully rational; if she is fully rational, she will be motivated to act in that way, and motivation is analysable in terms of desires and beliefs about means and ends. And, if a fully rational agent has an objective normative reason to act in some way, then if she acts in that way, her action is objectively desirable, inasmuch as all other agents would agree that it is desirable. In a later article, Smith supports this claim about the role of objective normative reasons, writing of such reasons that they are: 'propositions to the effect that this or that course of action is to some extent worth doing;…. the best account of such propositions is given by a dispositional theory of value' (2004, pp. 60-61).

Smith reads what it means for an action to be in some way right in a way that differs from my reading; he takes his account of objective normative reasons to give an account of right action,

but ignores the accessibility of those reasons to agents. His resolution of 'the moral problem' requires normative reasons to be connected to objective facts about what agents have reason to do, and he holds that those objective facts must abstract away from individual differences that arise from agents' practical and epistemic limitations. So, Smith's account of full rationality connects objective normative reasons to the potential consequences of acting in a certain way even when real human agents could not be aware of those consequences. He holds that the actions that are desirable, in other words, that have value, are those that fully rationally agents would take themselves to have reason to carry out.

Smith could not use my account of justifiable reasons in place of his account of objective normative reasons. An agent has justifiable reason to act in some way when she would take herself to have reason to act in that way if she reasoned in a way that was possible and appropriate. This means that agents can have justifiable reason to act in ways that fully rational agents would not take themselves to have reason to act. If, for example, someone with an irrepressibly gloomy disposition simply could not be brought to accept that she would not be doing her potential children a disservice by conceiving them, then no matter what a fully rational agent would take herself to have reason to do in the same set of circumstances, the agent has no justifiable reason to have children. So, if Smith were to use my account of justifiable reasons in place of his account of objective normative reasons, what has value would vary from person to person relative to their practical and epistemic abilities. This would mean that agents would not agree about what they had reason to do in a way that Smith holds would allow such reasons to be objective normative reasons.

Although Michael Smith and I reach different conclusions about what agents have reason to do, there is no tension between Smith's account of objective normative reasons and my account of justifiable reasons. Smith's account of objective normative reasons gives an account of what it means for an action to be ideally connected to the desirability of an action – fully rational agents would desire to carry out the action. In contrast, my account of justifiable reasons gives an account of what it means for an action to be one that agents ought to take themselves to have some reason to perform.

The relationship between idealisation, objective normative reasons and values developed by Smith is unique to him, but others who give accounts of objective normative reasons are similarly concerned to develop an account of objective normative reasons appropriately related to results and values, rather than an account of normative reasons that is appropriately related to actions that agents have some practical reason to carry out. Next, I explain this aspect of accounts of objective normative reasons given by a few major theorists who take significantly different approaches to ethics and value, namely, Tännsjö, Williams, McDowell and Schroeder.

Tännsjö, whose position was introduced above, develops an account of objective normative reasons that is directly focussed on the relationship between reasons and values (2010). Although he intends his account of reasons to hold independently of any particular moral theory, Tännsjö is a utilitarian who holds that agents have moral reasons to act in ways that

produce consequences that appropriately realise value, and he holds that all objective normative reasons are, in a sense, 'moral reasons'. He argues that it is unlikely that there are moral reasons that are also justifiable reasons; in other words, he holds that the shaded section of the 'moral reasons' set in the 'Normative reasons Venn diagram' is empty. Tännsjö's account of objective normative reasons is clearly developed to answer the 'What do agents *really* have reason to do?' question by developing a theory of which actions produce valuable outcomes rather than a theory about how real human agents have normative reason to act. Tännsjö's normative reasons get their normativity from moral principles. According to Tännsjö, I have a utilitarian moral reason to φ if φ-ing 'maximises the sum-total of well-being in the universe', and I have a Kantian moral reason to φ if, for example, φ-ing keeps a promise (2010, p. 28). If my action does not maximise the sum-total of well-being in the universe, or breaks a promise, then no matter how justified I thought I was in acting as I did, according to Tännsjö, I did not act for a moral reason.

Tännsjö claims that his position does not conflict with the 'ought implies can' requirement. He argues that an agent who has reason to bet on the winning horse in a race can always do so, because: 'We can bet on any one of them, so we can also bet on the winning one, even if we do not do so under this description' (Tännsjö, 2010, p. 33). In contrast, I take the claim that 'A ought to φ' to require that 'A can intentionally φ for a reason that would make φ-ing right'.

As argued in the last chapter, Bernard Williams' development of his account of internal reasons was probably partly motivated by his rejection of positions like Tännsjö's. Bernard Williams thinks that the dominant moral theories require agents to act in ways that grievously conflict with what matters to agents, so he argues that reasons must be tied to those things that can motivate agents to act (1981a, 1981b). He argues that any idealisation of normative reasons cannot assume that agents can have reason to do things that they would never be motivated to do (Williams, 1981a, 1995a, 1995b, 2001). Williams is concerned to show that agents who have nothing in their motivational sets that would, after sound deliberation, lead them to want to act in ways that realise what are typically taken to be moral values, have no reason to act in those ways (1981a, p. 110; 2001, pp. 92, 94). On Williams' account, moral values considered independently of agents' motivational sets, do not ground reasons for action. Nevertheless, Williams develops an account of objective normative reasons that is focussed on outcomes, or on values in the broad sense of the term, rather than on the action that it would be appropriate for a real human agent to carry out.

Williams argues that to find out what an agent has reason to do we need to imagine that the agent has carried out reasoning that idealises the set of things the agent could be motivated to do, while still limiting what she has reason to do to things that she could be brought to accept she has reason to do (1981a, pp. 102-103, 105, 108; 1995a, pp. 36, 38, 42). In §7.3, I quote Williams' acknowledgement that on his account of normative reasons agents will sometimes not be able to realise that they have normative reason to act in some way. Williams mentions 'unconscious obstacles' that might prevent an agent from working out that an action is one that she has reason to carry out (1995b, p. 188). Someone might, for example, be so scared of

bats that she cannot conceive of herself wanting to visit the Mulu caves in Sarawak. Yet even if the agent could never be brought to realise that she has reason to visit the caves, on Williams' account, she could still have an internal reason to visit the caves if she would want to visit the caves after working through her phobias; that is, if after sound deliberation she would accept that visiting the caves would be somewhat valuable for her, she has reason to visit the Mulu caves.

The way that Williams bases reasons on facts might also be taken as evidence that Williams is concerned that his account of reasons is properly related to actions that realise the right values rather than actions that agents would be justified in taking themselves to have reason to do. According to Williams, an agent has reason to φ when φ-ing will satisfy a desire that the agent would still have after the agent's subjective motivational set has been corrected by sound deliberation. This, on its own, would not entail that real human agents have reason to act in ways that they would not be justified in taking themselves to have reason to act. However, Williams writes:

> A member of [an agent's subjective motivational set], *D*, will not give *A* a reason for φ-ing if either the existence of *D* is dependent on false belief, or *A*'s belief in the relevance of φ-ing to the satisfaction of *D* is false. (Williams, 1981a, p. 103)

If *D* is a desire to drink from the glass in his hand, but the agent only wants to drink from the glass in his hand because he falsely believes it contains gin and tonic, then the above quote entails that the agent does not have reason to drink from the glass in his hand – not even if he is justified in believing that the glass contains gin and tonic. Williams claims that false beliefs cannot be the basis for reasons because if an agent's action is based on a false belief, the action will not satisfy the element of the agent's motivational set that would entail that acting in that particular way would produce the right result (2001, pp. 91-92). Given the link between Williams' theory and Humean accounts of reasons mentioned by Williams and others, this link between reasons and outcomes is unsurprising (Korsgaard, 1986, p. 8; Smith, 1995a, p. 118; Williams, 1981a, pp. 102, 104, 108).

John McDowell's account of reasons is also concerned with the value of the action (1995). McDowell outlines an account of objective normative reasons that uses a different form of idealisation from Williams and Tännsjö and is directed at realising a different form of value. John McDowell is a virtue ethicist and an externalist about reasons; he suggests that what an agent has objective normative reason to do is determined by what a perfectly virtuous individual, a '*phronimos*', would perceive there is reason to do (1979, 1995). A *phronimos* is not just someone who has the right kind of motivations, that is, someone who is motivated to be courageous, beneficent, and so on: a perfectly virtuous agent also perceives the world in a different way from those of us who are less ideal agents. If the virtuous agent and I were to describe a situation, her awareness of states and events would differ from mine. If the notion of a *phronimos* succeeds in the ways outlined by virtue ethicists, then the *phronimos* will have a sensitivity to objective values that the rest of us lack. Unlike Tännsjö's account of objective normative reasons, on a virtue ethicist's account of objective normative reasons such reasons will sometimes correspond to what the agent has justifiable reason to do. For

example, the *phronimos* might claim that if I cannot detect that the person is faking, I have objective normative reason to help someone who fakes a collapse next to me. Whether or not I can detect the fakery, my action could have value because it exhibits and develops virtue. But, the *phronimos*'s grounds for saying that I should help the faker can differ from the grounds for saying that I have justifiable reasons for helping the faker. Even if I correctly hold that I always have at least some normative reason to develop virtue in myself or others, it can be appropriate for me to fail to recognise that acting in a certain way would promote virtue.

Objective normative reasons are also the primary focus of Mark Schroeder's arguments in *Slaves of the Passions* (2008c, pp. 11-12).[102] Mark Schroeder develops a Humean account of reasons, which he calls 'Hypotheticalism' that aims to support a form of reductive realism about ethics (2008a, p. 178; 2008c). This in itself distinguishes his focus from mine, and shows his interest in debates about what justifies taking a certain end as conditioning of reasons. Schroeder says that he takes objective normative reason to be foundational, but says little about why he takes them to be foundational (2008c, p. 15). He knows that some Humean accounts of reasons have been criticised for being 'subjective' because they claim that desires are, or provide, reasons for action (Chang, 2006; cited in, Schroeder, 2008c, p. 21, fn 32). However, the subjectivity at issue in these criticisms arises from the role of desires within such accounts of reasons; it is not the subjectivity of Schroeder's 'subjective reasons', which were discussed in §6.1.2. Schroeder also mentions that some Humean theories of reasons are based on accounts of subjective reasons in a way that implies that this is a disadvantage (2008c, p. 13, fn 18). So, in offering a Humean account of reasons based on an account of objective normative reasons, Schroeder distinguishes his position from some criticised alternative positions. However, Schroeder writes that: 'the word "reason" in epistemology is typically understood to pick out its subjective normative sense, whereas in ethics it is usually objective reasons that are picked out with the word "reason"' (2008c, p. 113, fn 16). This quote suggests that Schroeder's real reason for treating objective normative reasons as foundational stems from his assumption that the common association of practical reasons with the achievement of goals or values is correct – or at least to be accepted for the purposes of his current work.

Schroeder aims to show that Humean accounts of reasons need not claim, or entail, that moral reasons are desires. He argues that the fact that someone has a desire is a background condition for someone having a reason to act in some way (Schroeder, 2008c, pp. 21, 23-40, 148-149). Schroeder gives the following account of what it is for something to be an objective normative reason:

> **Reason**   For *R* to be a reason for *X* to do *A* is for there to be some *p* such that
> *X* has a desire whose object is *p*, and the truth of *R* is part of what
> explains why *X*'s doing *A* promotes *p*. (Schroeder, 2008c, p. 59)

---

[102] Schroeder writes: 'The appropriateness of this terminology is an unfortunate consequence of etymology, as objective reasons … are no more *objective* than subjective reasons' (2008c, p. 12, fn 16, italics in original). Unfortunately, he does not explain this claim.

Where '*p*' and *X*'s desire, are Schroeder's 'background conditions', and made up of a combination of *END, VAL*, and *SIG*, and '*R*' is *STATE*, which I call the 'reason-explanans'. Schroeder illustrates his arguments using the example of Ronnie, who loves to dance, and has reason to go to a party because there will be dancing there (2008c, p. 1). Using the example of Ronnie to illustrate Reason: For the fact that there will be dancing at a party to be a reason for Ronnie to go the party, is for Ronnie to desire to dance, and for the fact that there is dancing at the party to be part of what explains why Ronnie's going to the party makes it likely that he will get to dance. Note that Schroeder specifies in his formulation of Reason that reason *R*'s truth must mean that an action will produce a certain outcome (the satisfaction of a desire). *R*'s truth does not on its own make *R* a reason. Assume that there will be dancing at the party, but that everyone who will be at the party thinks that Ronnie's dancing is so awful that if there is any indication that he intends to start dancing they will throw him off the balcony. The fact that there is dancing at this party will not explain why Ronnie's going to this party will put him in a position to dance; it won't. This requirement that 'the truth of *R* is part of what explains why *X*'s doing *A* promotes *p*' is a key difference between Schroeder's account of objective normative reasons and my account of justifiable reasons.

A second example of Schroeder's – Nate and the surprise party – provides a good illustration of the difference between the function of Schroeder's account of reasons and the function of justifiable reasons (2008c, pp. 33, 165-166). Nate loves surprise parties but hates learning about them beforehand. Schroeder writes that Nate has reason to go home because all his friends are waiting to throw him a surprise party. Nate cannot know that the surprise party is a reason to go home or use it to guide his actions, because if he knew about it, it would no longer be a surprise. Thus, there being a surprise party at his house can never give Nate a justifiable reason to go home; the interests of other people aside, it can never be appropriate for Nate to take the surprise party that people are about to throw him at his house to be a reason-explanans for him to go to home, because if he knows about the party it no longer gives him reason to go home. However, the fact that the surprise party cannot give Nate a justifiable reason to go home does not mean that the surprise party cannot be an objective normative reason-explanans for Nate to go home. As the case is described, Nate's predilection for surprise parties means that *STATE, END, VAL*, and *SIG*, favour Nate's going home. Why? Because if Nate goes home, his going home will produce an outcome with a certain value: Nate will get to enjoy his surprise party. So, on Schroeder's account of objective normative reasons, what agents really have reason to do is achieve certain outcomes.

It should be clear that none of the philosophers discussed above could use my account of justifiable reasons in place of their accounts of objective normative reasons. On my account of justifiable reasons:

> *A* has justifiable reason to *φ* in *C* if and only if, were *A*, in *C*, to reason in a
> way that is possible and appropriate, she would be aware of *STATE*, and take
> it that *END, VAL*, and *SIG*.

This account of justifiable reasons does not explain what has value; it merely includes it as a placeholder. As explained above and in §3.2, my use of 'value' is to be interpreted as broadly

as possible; I set no parameters or restrictions on what can count as 'valuable'. This is a gap in my account of reasons, and it also means that my account of justifiable reasons can contribute little to debates about the nature of value. My account puts limits on what can fill this role for justifiable reasons; whatever makes something valuable, for it to serve the role of *VAL* or *SIG* for justifiable reasons, it needs to be possible and appropriate for the relevant agent to take it to have the kind of value that would warrant acting on it.[103] An account of value that states that only actions that maximally improve overall wellbeing have value would be a useless addition to an account of justifiable reasons. The number of occasions when it would be empirically and practically appropriate for real human agents to take it that their actions will maximally improve overall wellbeing will be very small. And, if this account of value were used with my account of justifiable reasons, whenever it is not appropriate for agents to take it that an action will maximally improve overall wellbeing, the agents have no justifiable reason to act. However, this is a limit on what can count as valuable for justifiable reason relations, not a positive specification of what is of value. In contrast, philosophers who develop accounts of objective normative reasons develop the accounts they do partly because they intend to develop a clearer understanding of what it is for something to be of value. These philosophers tend not to argue for an account of value, but their theories of objective normative reasons are a step towards understanding what has value.

8.3 Objective normative reasons as guidelines

Accounts of objective normative reasons also serve a practical purpose. Justifiable reasons are not justified beliefs about objective normative reasons. In §6.3 I argued that:

> It is not the case that if *A* has a justifiable reason to *φ,* then she has a justified belief that she has an objective normative reason to *φ*.

> If *A* has a justified belief that she has an objective normative reason to *φ*, then she has a justifiable reason to *φ*.

So, if it is appropriate for an agent to have a justified belief that she has an objective normative reason to act in some way, the agent has a justifiable reason to act in that way. For example, an agent who justifiably believes that she is alone in a burning shed, would presumably be justified in believing that she has an objective normative reason to get out, and in such a situation, she would have a justifiable reason to leave. Debates about the nature of objective normative reasons help establish when an agent would be justified in believing that she has an objective normative reason to act in some way.

This attempt to identify when an agent would be justified in believing that he has good reason to act some way can be seen in the debate about internal and external reasons. Williams gives an example of a man who is unkind to his wife, and who, in spite of being remonstrated with, insists that he really sees no reason to be nicer to her. According to Williams:

---

[103] I argue in support of this claim in §7.1.3, and explain that this is not a claim that justifiable reasons must be internal reasons.

> There are many things [that can be said] about or to this man; that he is ungrateful, inconsiderate, hard, sexist, nasty, selfish, brutal, and many other disadvantageous things. I shall presumably say, whatever else I say, that it would be better if he were nicer to her. (Williams, 1995a, p. 39)[104]

But, Williams argues that the speaker may not say that the man's failure to be disposed to be nice to his wife results from irrationality – or say that the man has reasons to act that would never motivate him to act. If Williams is correct, then any time an agent has a justified belief that he could never be motivated to act in a certain way, the agent has a justified belief that he does not have an objective normative reason to act in that way. And, in such a case, he has no justifiable reason to act in that way.

Michael Smith's account of objective normative reasons illustrates a contrasting position on what it would be for an agent to have a justified belief that she has an objective normative reason to act. An agent may accept, as Smith suggests she should, that she has an objective normative reason to act in some way whenever she would be justified in believing that: she knows all the relevant facts; has no relevant false beliefs; and that her beliefs and desires about the circumstances are coherent. Of course, this does not presuppose that she knows anything about Smith or his meta-ethical theories, just that she shares his conception of what it is to have a normative reason to act.[105] For example, an urban search and rescue (USAR) team leader could have a justified belief that she has a (Smithian) objective normative reason to call for USAR dogs because she has a justified belief that she knows all the relevant facts and has no false beliefs about the circumstances, and a justified belief that her beliefs and desires about the search and rescue task are coherent. If this is the case, then she can also be said to have a justifiable reason to call for USAR dogs. Of course, the agent could be wrong about what she has objective normative reason to do. She may be justified in believing that she knows all the relevant facts but still be missing some crucial piece of information. Nevertheless, she can use Smith's theory as a guide to right action. An agent whose conception of objective normative reasons corresponds to Smith's, but who is unsure of the facts that relate to her situation, unsure of the relevance of information to her situation, or

---

[104] See also, (Williams, 1981a, p. 110).

[105] Michael Smith seems to hold that his account of objective normative reasons corresponds to the concept of normative reasons that is generally accepted by agents. Smith responds to Geoffrey Sayre-McCord's suggestion that agents can understand and act on normative reasons without possessing Smith's concept of objective normative reasons as follows (Sayre-McCord, 1997, p. 81):

> Those who are competent with the concept of a normative reason really are sensitive in their application of the concept to the way in which failures of information can undermine normative reason claims; they really are sensitive in their application of the concept to the way in which the unavailability of a certain sort of ideal justification—a lack of coherence and unity—can undermine normative reason claims; and they really are sensitive in their application of the concept to the way in which a failure to be motivated in accordance with allegedly accepted normative reason claims, at least absent practical irrationality, can undermine the genuineness of the acceptance of the normative reasons claim. Those who fail to exhibit these sensitivities are not properly competent in their use of the concept of a normative reason. What justifies our attribution to them of the complex concept of what they would desire if they had a maximally informed and coherent and unified set of desires is thus that their possession of this concept best explains these discriminative abilities that they manifest in their judgments and inferences. (Smith, 1997a, p. 105)

who knows that her set of beliefs and desires about her situation is incoherent, should not take herself to be justified in believing that she has an objective normative reason to act in a certain way. However she could still have a justifiable reason for acting in that way. Perhaps the USAR team leader knows that she is missing crucial information about the dangers of contacting the USAR dog team. Nevertheless, if it is appropriate for her to take the situation to warrant calling in the dog team, she has a justifiable reason to do so.

Objective normative reasons can serve as a guide to what agents have justifiable reason to do in a second way. My account of justifiable reasons gives value a key role within reason relations, but makes no effort to explain what makes an action or the outcome of an action valuable (other than its supposition that there is something valuable about agents appropriately taking account of justifiable reasons). Philosophers and other human agents want to know what it is that they have reason to try to achieve, both for major achievements like dealing with famines and floods, and for actions that it seems odd to call 'achievements', such as sorting out ablutions. If an account of objective normative reasons were to succeed in explaining what has value and when values significantly trump each other, this would alter what agents have justifiable reason to do, and alter it even for situations where it would be inappropriate for them to form a justified belief that they have an objective normative reason to act in some way.

8.4 Conclusion

Objective normative reasons and justifiable reasons serve different functions. As I argued in Chapter 5, they do not compete because they are not in the same race. Objective normative reason relations and justifiable reason relations offer different explanations of why an agent has normative reason to act in a particular way. Justifiable reasons are reasons that it is appropriate for agents to take themselves to have at least some reason to act on. Objective normative reasons are indicators that an agent's acting in some way will yield some value.

The relationship between objective normative reasons and value leads many objective normative reasons theorists to take the normativity associated with objective normative reasons to be derivative on the normativity of the associated values. This means that many objective normative reasons theorists take the accessibility of objective normative reasons to be less important than the likelihood that acting on such reasons will, in fact, yield the relevant value. Torbjörn Tännsjö is an extreme example; if people can never know which actions will realise the most value, then on Tännsjö's account of objective normative reasons, they can never act on normative reasons (2010, pp. 32, 33-34, 90-91). Jonathan Dancy is an exception to this rule, as he claims that only states of affairs that are accessible to agents can be objective normative reasons (2000, pp. 56-59, 65-66; 2004, pp. 158-159). But for Dancy, the states of affairs that are objective normative reasons are only those that indicate that acting in a certain way will realise appropriate values (2000, p. 29).

The relationship between objective normative reasons and values gives objective normative reasons an ability to serve a purpose that justifiable reasons cannot serve. My brief exposition of Michael Smith's argument in *The Moral Problem* showed one use of objective normative reasons to resolve problems in meta-ethics, and I argued that justifiable reasons could not serve this purpose. I gave even briefer explanations of the roles that objective normative reasons play in other theorists' research programmes that could not be played by justifiable reasons. I argue throughout this thesis that it is worth developing an account of the normative reasons that real agents should take themselves to have some reason to act on, in other words, an account of justifiable reasons. The importance of justifiable reasons does not, however, reduce the importance of objective normative reasons, or the value of debates about the correct account of objective normative reasons.

Although justifiable reasons and objective normative reasons serve different functions, as stated in §4.4, the objective normative reason-explanans and justifiable reason-explanans identified by *STATE* are metaphysically the same kind thing: both are state of affairs. This means that the same state of affairs can be an objective normative and justifiable reason-explanans for acting in a certain way. What differs for these different forms of reason relation are the things that pick out certain states of affairs as reason-explanans. While the *STATE* picked out as a reason-explanans for objective normative reasons will be derived from an account of normative reasons designed to identify the actions that will realise certain values, the *STATE* picked out as a reason-explanans for justifiable reasons will be derived from an account of normative reasons designed to identify the actions that agents would be justified in taking themselves to have some reason to carry out.

Objective normative reasons cannot be used to explain rationality because they function as indicators of value rather than as indicators of actions that it would be in some way right for agents to carry out. However, this is a function that justifiable reasons can serve. This is the topic of the next chapter.

9. Rationality and Reasons

9.0 Introduction

Acting rationally amounts to acting appropriately in response to justifiable reasons, or so I will argue.[106] This relationship between rationality and normative reasons is one of the ways in which my account of justifiable reasons can be distinguished from objective normative reasons.[107] Two commonly mentioned analyses of rationality resemble the approach I take. Rationality is sometimes said to amount to acting appropriately in response to reasons or to involve acting appropriately in response to beliefs about reasons. Objections have been raised to both of these analyses of rationality. I argue that my account of rationality does not succumb to these objections. A more significant problem for both my account of rationality and my account of justifiable reasons arises from arguments that claim to show that rationality is a relationship between attitudes rather than a relationship between the world and actions.

Like the word 'reason', the word 'rational' is used to refer to a number of different concepts. In §9.1, different uses of the term are described and I explain in general terms what I do and do not take the word to mean. In §9.2, I briefly explain my claim that agents act rationally if and only if they respond appropriately to justifiable reasons.

In §9.3 and §9.4 I examine two common approaches to analysing reasons in terms of rationality. Rationality is sometimes said to involve appropriately responding to reasons, where 'reasons' means 'normative reasons'. This approach to analysing reasons is usually rejected. §9.3 shows why objections to this claim that rationality involves appropriately responding to reasons do not affect my claim that rationally is equivalent to responding appropriately to justifiable reasons. In §9.4, I examine the claim that agents act rationally when they respond appropriately to their beliefs about what they have reason to do. This claim about rationality could be taken to mean one of two things. In §9.4.1, this is taken to be a straightforward claim that rationality requires agents to respond appropriately to their beliefs about what they have normative reason to do. In spite of the initial plausibility of this analysis of rationality, it is usually taken to fail because of the so-called 'bootstrapping objection'. So, in this section I explain this objection, explain the response made to this objection by some objective normative reasons theorists, and explain why the bootstrapping objection does not affect my analysis of rationality in terms of justifiable reasons. In §9.4.2, I consider the claim that rationality requires agents to respond appropriately to their justified beliefs about what they have reason to do. This amounts to a claim that rationality requires appropriate response to certain forms of subjective normative reasons. I argue that this account of rationality is partly correct, because responding appropriately to justified beliefs

---

[106] Errol Lord, whose position was outlined in §4.4, argues for a similar claim (2010, pp. 293-294).
[107] See §4.1.

about reasons is sufficient for acting rationally, but I do not accept that this is necessary for acting rationally. Agents can be unaware of the reasons they have for acting, either due to ignorance or because they can act for reasons of which they are not consciously aware. So, the claim that rationality requires agents to act appropriately in response to justifiable reasons succeeds in a way that explains the limited success of the claim that rationality requires agents to respond appropriately to their justified beliefs about what they have reason to do.

Some people argue that rationality is a relationship between cognitive attitudes such as beliefs, desires and intentions. But, if rationality is a relationship between cognitive attitudes, either my account of rationality or my account of justifiable reasons must be flawed; rationality either cannot be explained in terms of justifiable reasons or justifiable reasons must be cognitive attitudes rather than states of affairs. This objection is explained in §9.5, and I outline two possible responses to the objection. First, in §9.5.1, I briefly consider whether epistemological externalists can save my argument. Next, in §9.5.2, I discuss the consequences for justifiable reasons and objective normative reasons of accepting epistemological internalism. I conclude that if this theory is correct, many theories of objective normative reasons would also have to alter their claims about the nature of reason-explanans.

The issues discussed in this chapter relate to a number of important issues that fall outside the scope of this thesis. Arguments about the relationship between rationality and reasons are connected to arguments about whether normativity is only a property of reasons, and arguments about whether one ought to be rational. These arguments are not considered here. But, if acting rationally is equivalent to acting appropriately in response to justifiable reasons, this supports those who argue that normativity is always grounded in reasons, and opens another route to arguing that agents ought to be rational.

9.1 Forms of rationality

Like the word 'reason', the word 'rational' is used to refer to a number of different concepts that are distinguished using technical terms. Niko Kolodny distinguishes between 'subjective rationality' and 'objective rationality' (2005, pp. 209-210). Although Kolodny does not express it this way, 'objective rationality' requires agents to act as they have overall objective normative reason to act. So, to be objectively rational an agent must act as some particular gods'-eye view of the world would recommend he acts. The term 'substantive rationality' is often taken to refer to an agent's doing what he has most reason to do, which suggests that on one use of the term, 'substantive rationality' corresponds to Kolodny's objective rationality. [108] Kolodny takes 'subjective rationality' to require consistency between an agent's intended ends and the means the agent takes to be available to bring about those ends (2005, pp. 209-210). Thus, Kolodny's 'subjective rationality' seems to correspond to 'means-

---

[108] Dancy mentions this in 'Reasons and Rationality' (2009, pp. 95-96). Errol Lord and Amelie Rorty, use 'substantive rationality' in very different ways (Lord, 2010; Rorty, 2009). Lord's use of 'substantive rationality' resembles my account of 'rationality'.

ends calculative rationality', that is, 'maximizing benefits with minimal costs towards a pre-determined end, independently of the content of that end' (Rorty, 2009, p. 350). This means-end way of understanding what it is to be rational is somewhat controversial. If James wants to inherit Uncle John's money more than he wants anything else, there might be a sense in which it is irrational for James not to push his Uncle John out of a boat, but it seems odd to call drowning Uncle John 'rational'.[109] Pushing uncles out of boats is probably not 'objectively rational'. I argue later that while it might be irrational for agents like James not to act as they insist they currently most want to act, this does not entail that acting as they currently most want to act is rational. As I use the term, 'rationality' requires more than consistency between what an agent currently intends to achieve and the agent's taking the means to that end, but less than acting as someone with a gods'-eye perspective on the world would act.[110]

Jonathan Dancy suggests three possible contenders for the title of 'requirement of rationality':

1. Do what you have most reason to do.

2. Do what, if things were as you suppose them to be, you would have most reason to do.

3. Do what you believe yourself to have most reason to do. (Dancy, 2009, p. 95)

Dancy claims that '1' is taken by others to involve 'substantive rationality', but Dancy takes it to be a requirement of reason. '2' is taken by others, but not Dancy, to be a 'requirement of rationality'. And '3' is what Dancy takes to be a 'requirement of rationality' (2009, pp. 95-96). None of Dancy's rational requirements correspond to my account of rationality.

There are two situations where I would be happy to accept the first proposition, that is, the claim that rationality requires agents to 'Do what you have most reason to do.' First, there seems nothing wrong with calling this approach to rationality an account of 'substantive rationality'. But, this is not a form of rationality that is of interest in this thesis. Second, I would accept the claim that agents act rationally when they act as they have most reason to act if 'reason' were taken to refer to 'justifiable reason'. However, Dancy uses 'reason' to refer to 'objective normative reason'. Dancy prefers not to use 'rational' to refer to situations where agents do what they have most reason to do. Dancy wants to use 'rational' to refer to situations where agents are not 'at odds with themselves', and he accepts that agents can justifiably fail to know what they have objective normative reason to do (2009, pp. 96, 97).

I am also not interested here in the rationality that you would exhibit if you were to 'do what, if things were as you suppose them to be, you would have most reason to do'. This requires both too little and too much of agents. It requires too little of agents to correspond with the form of rationality in which I am interested, because an agent's understanding of her circumstances can be unjustified. She may, for example, have been careless when she came to her understanding of her circumstances. This way of understanding rationality requires too

---

[109] Example from R. M. Hare (1972).

[110] I expand on this claim in §9.4.1.

much of agents because even an agent whose understanding of a situation includes all relevant true beliefs and no relevant false beliefs, may not be in a position to use her reason to what she has objective normative reason to do. Dancy's reason for rejecting this way of understanding rationality differs from mine. Dancy thinks that it is appropriate to praise someone who does what she would have most reason to do if things were as she supposed, but does not think it is appropriate to think of her as acting rationally. He sees this as expressing a relationship between 'reasons of a special sort…. [that favour] combinations of belief and action] (2009, p. 99). This approach to rationality is discussed more below in §9.4.

Dancy's preferred example of a rational constraint is his third suggestion: 'Do what you believe yourself to have most reason to do'. Notice that this statement is not about objective normative reasons at all. It is about agents' beliefs about reasons. This subjective reasons approach to rationality is too weak to correspond to the form of rationality that interests me. As I argued in Chapter 6, agents can believe they have most reason to act in certain ways for bad reasons. This way of thinking about rationality requires agents not to be 'at odds with themselves', but only on a very narrow way of thinking about what it means to be at odds with oneself (Dancy, 2009, p. 96). The difference between my account of rationality and that of Dancy becomes clearer later.

My analysis of rationality also doesn't neatly fit into the categories of 'procedural' or 'substantive' rationality – at least as Derek Parfit defines them. According to Parfit, 'procedural rationality' involves appropriately following rules of reasoning, while 'substantive rationality' requires that we have the right kinds of desires, values, or intentions (1997, pp. 99, 101, 102, 116).[111] So, James might be said to be procedurally rational if he carries out the reasoning necessary for him to work out how to kill Uncle John, but substantively irrational because he intends to kill him. Joseph Raz rejects the claim that we can distinguish between substantive and procedural rationality, arguing that an agent could not be procedurally rational while being substantively irrational, or vice versa (Raz, 1999, pp. 72-73). Although I disagree with Raz's claim that we can't make sense of the distinction, I think he has good reasons for questioning the idea that an agent could be just procedurally or just substantively rational. Like Raz, I take 'rationality' to involve getting both reasoning and mental states in order. What I mean by this becomes clearer below.

When people claim that acting rationally involves acting appropriately in response to reasons or beliefs about reasons, they mean that acting rationally involves acting appropriately *in response to* those reasons or beliefs. If James accidentally pushes his uncle overboard, he is not acting appropriately in response to reasons. For James's action to exhibit even means-end rationality, James needs to push Uncle John overboard as part of his effort to kill his uncle. But this does not mean that 'acting in response to reasons' requires the agent to be

---

[111] People sometimes talk as though there are some standards for rationality that are external to agents' abilities, but that do not fit with the usual use of the term 'substantive rationality'. For example, certain kinds of insanity are thought to result in irrational beliefs or actions. I would argue that these beliefs and actions are called 'irrational' because they fail to show a certain kind, or degree, of responsiveness to normative reasons that is expected of autonomous agents.

consciously aware of the reasons that he is responding to. As long as the action is appropriately connected to the reason, agents can respond appropriately to reasons without being consciously aware of the reasons to which they are responding.[112] Thus, if James is an experienced sailor and senses that the boat will overturn if he pushes Uncle John starboard, he may be acting in response to reasons when, unreflectively, he positions himself so that he can push his uncle off the port side. James's movement results from a well-developed knowledge of boats and balance. Arguably, agents can act rationally even when they have never decided for themselves that they will form a certain habit. Someone trained as a child to smile and say 'thank you' whenever something is offered may never have thought deeply about the reasons for acting in that way, yet she may still be acting rationally when she acts in that way.

Raz takes the claim that agents can be acting in response to reasons when they act unreflectively to an extreme. According to Raz, rational action must be 'automatic, rather than a product of deliberation and decision' (1999, p. 72). To support this claim, Raz argues that an agent with obsessive-compulsive disorder is not rational if she has to make a deliberate, reflective decision not to check that her door is locked after repeatedly checking that it is locked. Although I accept that rational actions can be automatic, I disagree with this claim that rationality never involves deliberation and decision. The agent with obsessive-compulsive disorder understands that there is no objective need to check her door again, understands that for her mental health she needs not to check her door again, and acts appropriately. Sometimes setting time aside to think about precisely how one has reason to act is a necessary precursor to acting rationally.

As explained in §2.1, although I usually write as though the end point of practical reasoning is an action, the end point of practical reasoning can be taken to be an agent's forming an intention to act. An agent may have normative reason to form an intention to act in a situation where an internal or external obstacle prevents her from acting. In such a situation, the agent acts in response to reasons when she forms the intention. As long as she genuinely intends to act, her failure to carry out the act is irrelevant.

When people discuss claims about the relationship between reasons and rationality, they commonly suggest that acting rationally involves acting appropriately in response to overall reasons rather than pro tanto reasons. Accept, for the sake of the example, that James's lust for money justifies killing Uncle John. If either stabbing Uncle John or pushing him overboard will kill him, but the latter is a better way to conceal the murder, then James may have pro tanto reason to stab Uncle John, but overall reason to push him overboard. In this situation, it is means-end rational for James to push his uncle overboard and not means-end rational for James to stab him. In spite of the plausibility of restricting claims about the relationship between rationality and reasons to claims about rationality and overall reasons, this restriction is unnecessary. If James evaluates the options and pushes his uncle overboard, he is acting appropriately in response to both his overall and pro tanto reasons. The

---

[112] I argue for this in §3.1.

appropriate response to a pro tanto reason that conflicts with an overall reason is to not act for that reason. However, it remains the case that the reasons that it is rational for agents to actually act on are the reasons that the agents have overall reason to act on.

John Broome takes a slightly different approach; rather than expressing the reasons-rationality relation in terms of overall reasons, he argues that responding appropriately to reasons is best taken to mean that 'you must *F* whenever your reasons require you to *F*' (2007a, p. 350). This is compatible with my claim about responding correctly to pro tanto reasons that are not also overall reasons, because it means that not acting in some way when your reasons do not require you to act in that way counts as responding correctly to reasons. But, this way of expressing the relationship seems to exclude certain appropriate responses to reasons from being appropriate responses to reasons. An agent's reasons may mean that acting in some way is an okay thing to do, rather than a required action. If, on a lazy Sunday morning, some reasons favour staying in bed and reading, while others favour going out for brunch, an agent could be responding appropriately to reasons whichever option she chooses. In such a situation, the agent seems not to be 'required' to act in either way. The agent might be said to have overall reason to either read or go out for brunch, but it still seems too strong to say that she is required to either read or go out for bunch. *F*-ing when your reasons require you to *F* is just one way of responding appropriately to reasons.

9.2 Rationality and justifiable reasons

What does the claim that acting rationally is equivalent to responding appropriately to justifiable reasons amount to? Recall that an agent has a justifiable reason to act in some way when she would take herself to have reason to act in that way if she reasoned in a way that is possible and appropriate given her circumstances. (I will omit the word 'possible' in the rest of this chapter, and suppose that if something is appropriate, it is possible.) I argued in §4.2 that it is appropriate for agents to take themselves to have reason to act in some way when it is appropriate for them to be aware of some *STATE*, and for them to hold that *END, VAL*, and *SIG* are all the case. So I claim that agents act rationally if and only if they respond appropriately to the sets of *STATE, END, VAL*, and *SIG* that it is appropriate for them to take to be the case. The form of rationality that arises from this definition requires more than means-end rationality; it avoids problems associated with mere means-end rationality without requiring agents to be objectively rational.

Consider an example. You are ill and your doctor recommends that you take the red pills. If you reason appropriately given your circumstances, you will conclude that if you do not take the pills, you will probably stay ill and unhappy. You have no reason to believe that taking the pills will have any unpleasant consequences for anyone. In this example, you have justifiable reason to take the medicine because it is appropriate for you to be aware that:

(1) your doctor told you to take the red pills.

And, it is appropriate for you to hold that:

(2) your doctor is the best guide to what you need to do to become well again;

(3) getting well has value; and,

(4) following the doctor's advice will not conflict so much with other values that the conflict prevents you from having reason to act on your doctor's advice.

(1)	*STATE*	A state of affairs.	Your doctor tells you to take the red pills.
(2)	*END*	Given the state of affairs, an action will (probably) yield an effect.	Your doctor's telling you to take the red pills is a sign that taking the red pills might improve your functioning.
(3)	*VAL*	The effect (probably) has positive value.	Improving your functioning has some value.
(4)	*SIG*	The positive value of the action (probably) remains significant in spite of any negative values produced by the action.	Taking the red pills does not conflict so much with other values that it would be nonsense to say there is reason to do it.

Assume that you have overall justifiable reason to swallow the pills. In situations like this, I claim that swallowing the red pills is an appropriate response to justifiable reasons, so it is rational for you to swallow them.

A programme that analyses rationality in terms of reasons could be claiming one of two things.[113] First, the claim might be that rationality is equivalent to responding appropriately to reasons or beliefs about reasons. To establish this claim you need to show that responding appropriately to reasons is both necessary and sufficient for being rational. Second, the claim might be that rationality can be reduced to responding appropriately to reasons or beliefs about reasons. As Broome points out, rationality is only reducible to reasons if the first equivalence claim is correct (2007a, pp. 349-350). I argue that rationality is equivalent to responding appropriately to justifiable reasons, but do not argue that rationality can be reduced to responding appropriately to justifiable reasons.

In the following sections, I examine two relationships between reasons and rationality that resemble mine. I also examine objections to those proposed relationships, and consider whether those objections also impact on my claims about the relationships between justifiable reasons and rationality. First, I consider the claim that rationality cannot be equivalent to responding appropriately to normative reasons because reasons and rationality can conflict. Second, I consider the claim that rationality requires agents to respond appropriately to beliefs about reasons.

---

[113] Broome introduces this distinction (2007a, pp. 349-350).

9.3 Rationality as appropriate response to reasons

The analysis of the relationship between reasons and rationality that seems most similar to mine is the analysis of rationality as an appropriate response to reasons. Joseph Raz, for example, argues that practical and theoretical rationality involve appropriate responses to reason and reasons (1999). Exactly what Raz means by 'reasons' is unclear, because he intentionally leaves the term 'reason' ambiguous (1999, p. 55). Despite the vagueness of claims that rationality involves responding appropriately to reasons, a number of philosophers have argued that rationality cannot be equivalent to appropriately responding to normative reasons.

Philosophers who claim that rationality cannot amount to responding appropriately to reasons support their claim by arguing that what rationality requires can differ from what an agent has normative reason to do. This argument is given by people like Broome and Kolodny, who take normative reasons to be what I have called 'objective normative reasons', and if 'normative reasons' refers to objective normative reasons, their argument is correct (Broome, 2007a, pp. 352-353; Kolodny, 2005, p. 512). If acting rationally amounted to appropriately responding to objective normative reasons, then there could be a conflict between reasons and rationality. However, this objection doesn't refute my claim that rationality involves responding appropriately to justifiable reasons.

Imagine, again, that your doctor has recommended that you take the red pills. The fact that your usual, reliable doctor has prescribed this medicine for you means that it is rational for you to take the medicine. But, assume that your doctor got it wrong: you have an unknown fatal allergy to that medicine. Objective normative reasons theorists will argue that even though you have no objective normative reason to take the medicine it is still rational for you to take the medicine, because no one involved could be expected to know about your allergy.[114] Thus, for objective normative reasons theorists, reasons and rationality come apart; there is a conflict between what it is rational for you to do (swallow the red pills) and what you have objective normative reason to do (not swallow the red pills).

This claim that there can be a conflict between what it is rational for an agent to do and what the agent has normative reason to do does not apply to justifiable reasons. Justifiable reasons are reasons that it is appropriate to expect agents to realise they have reason to act on. When objective normative reasons theorists argue that normative reasons can conflict with rationality, they usually describe situations where it is not appropriate to expect agents to take themselves to have reason to act in response to the relevant objective normative reasons. John Broome gives the following example:

> The fish on the plate in front of you contains salmonella…. [All] your reasons
> together require you not to eat it. But you have no evidence that the fish contains

---

[114] Dancy might not claim this because he might hold that his epistemic filter means that your allergy to the drug cannot serve as a reason for your doctor not to prescribe the drug or for you not to take the drug (2000, pp. 56-59, 65-66). See §5.3 for more details.

salmonella. Then you might eat it even though your reasons require you not to, and nevertheless you might be rational. (Broome, 2007a, p. 352)

I agree with Broome that it is rational for you to eat the fish, and also agree that you have no objective normative reason to eat the fish. However, as long as it is not appropriate to expect you to have evidence that the fish contains salmonella, you have a justifiable reason to eat the fish. So, although your objective normative reasons conflict with what you have reason to do, your justifiable reasons coincide with what it is rational for you to do.

Rather than introducing the notion of justifiable reasons in response to these problems with explaining rationality in terms of objective normative reasons, the standard move at this point is to consider the possibility that rationality involves responding appropriately to beliefs about reasons. However, there are also problems with this analysis.

9.4 Rationality and beliefs about reasons

One alternative to claiming that acting rationally involves responding appropriately to objective normative reasons is to claim that agents act rationally when they respond appropriately to their beliefs about normative reasons. Given common analyses of subjective normative reasons, this approach is similar to claiming that rationality involves appropriately responding to subjective normative reasons. [115] So, whenever I write 'acting rationally involves responding appropriately to beliefs about normative reasons' below, this could be exchanged for 'acting rationally involves responding appropriately to subjective normative reasons'.

The ambiguities that affect other work on practical reason also affect arguments about the relationship between rationality and beliefs about reasons. The claim could be that rationality involves appropriate response to beliefs about objective normative reasons or beliefs about normative reasons of any kind. If so, acting rationally just is acting as you believe you have overall reason to act. Alternatively, rationality could involve appropriate response to justified beliefs about objective normative reasons or justified beliefs about normative reasons of any kind. If so, acting rationally requires agents to act in response to justified beliefs about what they have overall reason to do. When the implications of the claim differ depending on whether it is normative reasons of any kind or objective normative reasons that are at issue, I state which form of reasons is at issue. Ambiguity is also an issue in the reference to 'reasons' in the sentence 'rationality involves an appropriate response to beliefs about reasons'. I read this as a reference to beliefs about what I call 'reason relations', not as a reference to beliefs about what I call 'reason-explanans'. To avoid the clumsy terms 'reason relations' and 'reason-explanans', I make the reference to reason relations clear by writing that 'agents are rational when they act appropriately given their beliefs about *what they have overall reason to do*'. I consider claims that rationality involves acting in response to beliefs about reason-explanans later.

---

[115] See Chapter 6.

9.4.1 Rationality requires an appropriate response to beliefs about reasons

The claim that agents act rationally when they act appropriately given their beliefs about what they have overall reason to do seems plausible, but philosophers have argued that it is false. It seems plausible because agents are usually expected to act as they believe they have most reason to act. Acting as you believe you have overall reason to act involves one variety of means-end rationality. Assume, for example, that you believe that you have overall reason to take the blue pills rather than the red ones. If you feel ill, want to get better, believe the blue pills would cure you, and yet decide not to take them, then, in one sense of the word, your actions seem to be irrational. If you believe that you have most reason to act in some way and fail to form an intention to act in that way, your set of beliefs and intentions is incoherent, and this incoherence is sometimes called 'irrationality'. Yet the idea that rationality requires agents to act appropriately in response to their beliefs about what they have reason to do has been criticised for involving what is called 'bootstrapping'.

According to the bootstrapping argument, if rationality requires you to act in whatever ways you believe you have overall reason to act, and you assume that there is a reason to do what rationality requires you to do, just believing you have overall reason to do something gives you reason to do it (Bratman, 1987, pp. 24-27; Broome, 2001, pp. 98-99; Kolodny, 2005, pp. 512, 514-515). The bootstrapping argument goes as follows:

> P1: Assume (for a *reductio*) that rationality requires agents to act as they believe they have overall normative reason to act.
>
> P2: *A* believes she has overall normative reason to φ.
>
> P3: Agents should act rationally.
>
> C1: Agents should act as they believe they have overall normative reason to act.
>
> C2: *A* should φ.
>
> P4: An agent ought to φ if and only if the agent has overall normative reason to φ.
>
> C3: *A* has overall normative reason to φ.

If this argument is sound, agents have overall normative reason to do whatever they believe they have overall normative reason to do. But, this is just false. People sometimes believe silly things for silly reasons. You might believe you have overall normative reason to take the blue pills because blue is your favourite colour, but it does not follow from this that you have overall reason to take the blue pills. Any argument must be fallacious if it purports to show that whenever an agent believes she has overall reason to act in some way, she has overall normative reason to act in that way.

The fourth premise in the bootstrapping argument claims that agents ought to act as they have overall normative reason to act. I have assumed that this sentence is meant to be read as referring to 'overall objective normative reasons'. However, if one accepts the adage that 'ought implies can', it does not make sense to claim that agents ought to do what they have overall objective normative reason to do. Agents often cannot know what they have objective normative reason to do, so they cannot base their actions on what they have objective normative reason to do. This being so, it cannot be the case that agents ought to act on their objective normative reasons. I don't think that this matters much to the argument, even though removing this premise gets rid of the egregious 'bootstrapping'. Concluding that an agent ought to act in some way because she believes she has overall reason to act in that way still leaves the argument looking questionable, because her belief that she has overall reason to act in that way might be unjustified.[116]

Objective normative reasons theorists sometimes introduce the distinction between wide and narrow scope oughts to try to avoid this bootstrapping problem.[117] Means-end rationality seems to require that you ought to, if you believe you have overall reason to take the blue pills, take the blue pills. However, this does not mean that you ought to take the blue pills. It means that you ought either decide that you do not have overall reason to take the blue pills or take the blue pills. Broome suggests that treating the 'ought' as having a wide scope avoids the bootstrapping problem (2007a, 2007b). However, Kolodny argues that Broome's response does not remove the problem because not all oughts are wide-scope oughts (Kolodny, 2005, pp. 514-542). Dancy argues that this approach misses something (2009).[118] On this way of thinking about reasons, the rational thing to do, if you believe you have most reason to take the blue pills, is either decide you do not have overall reason to take the blue pills or take the blue pills. Assume that you have a philosophy degree and you accept Broome's claims. You decide to toss a coin to decide what to do. Heads you stop believing you have overall reason to take the blue pills, tails you take them. As Dancy argues, there seems to be something wrong with the claim that you are rational if you act in whichever way the coin toss advises you to act.

I agree that agents can respond appropriately to their beliefs about what they have objective normative reason to do without acting rationally. This proposed relationship between beliefs about reasons and rationality leaves out information about agents and reasons that is needed

---

[116] There is a second argument for rejecting the claim that agents ought to do what they have overall normative reason to do. The claim that an agent has overall reason to φ might mean that the agent has more reason to φ than act in any other way, even when the agent's situation is such that saying she 'ought' to do anything is too strong a claim. However, this would not be a problem if 'oughts' come in a variety of strengths. Saying 'you ought to φ', might sometimes amount to nothing more than 'it would be a bit daft for you not to φ'. As in, 'there is no requirement for you to go out for brunch rather than read in bed, but you are a bit daft staying in bed when your pleasure is the only thing at stake and you would get more pleasure from going out for brunch'.

[117] See, for example, (Broome, 2007b; Dancy, 2009; Kolodny, 2005). Broome's position is discussed in more detail in §9.4.2.

[118] Dancy uses this argument in a different context, when he discusses the claim that rationality requires agents to '[do] what, if things were as you suppose them to be, you would have most reason to do' (2009, pp. 95, quote, 102-103).

for any assessment of what it would be rational for an agent to do. Return to the example where your doctor has prescribed you medicine. This time assume that your doctor prescribed you and your partner different medicines. You have forgotten whether you are supposed to take the blue pills or the red ones. You cannot read the labels on the bottles because your glasses are upstairs, but you believe that the blue pills are probably yours, and conclude that you have overall reason to take them. Is it rational for you to swallow the blue pills? When the circumstances are explained in this way, it seems irrational for you to take the blue pills in spite of your belief that you have overall reason to do so. If you swallow the blue pills and the blue pills are yours, the success of your action was down to luck, not the rationality of your action. Under this brief description of the circumstances, it seems that the rational thing for you to do is get your glasses and check the names on the bottles. If you have doubts about which pills will make you well, you have a justifiable reason to check which pills are the right ones. This means that I hold that an agent can act appropriately in response to her beliefs about what she has overall normative reason to do and yet act irrationally if her belief about what she has overall normative reason to do is unjustified. So, although my account of normative reasons allows states of affairs to count as reasons when they are not objective normative reasons, my account of rationality requires more than mere means-end rationality.

Examples of agents failing to take what they believe to be the means to their currently accepted ends are examples of irrationality, but they are not a sign that agents act rationally when they act according to their current beliefs about the best means of achieving their currently desired ends. Messing up the relationship between means, ends and actions is one way to fail to act in a justifiable way. If you do so, you are arguably irrational. (Peter Railton argues that this incoherence should be understood as agent-incompetence, that is, as a failure of practical competence, rather than as a failure of rationality (2009, pp. 113-114).) However, getting the means-end relation right is necessary, but not sufficient for being rational. Imagine, again, an agent who is unsure whether the blue pills or the red pills are hers, but who believes that she has overall reason to swallow the blue pills. In such circumstances, she would be means-end irrational if she did not swallow the blue pills. However, because, given her circumstances, her belief is not sufficient to justify her action, it would not be rational for her to swallow the blue pills.

My claim that acting rationally involves acting appropriately in response to justifiable reasons does not succumb to the bootstrapping problem. To see this, replace 'A believes that' with 'It is appropriate for A to hold that' in the bootstrapping argument:

> P1: Assume that rationality requires agents to φ when it is appropriate for them to hold that they have overall normative reason to φ.
>
> P2: It is appropriate for A to hold that she has overall normative reason to φ.
>
> P3: Agents should act rationally.
>
> C1: Agents should φ when it is appropriate for them to hold that they have overall normative reason to φ.
>
> C2: A should φ.

P4: An agent ought to φ if and only if the agent has overall justifiable reason to φ.

C3: *A* has overall justifiable reason to φ.

The bootstrapping argument does not work for justifiable reasons because what agents believe they have reason to do has no direct connection to what agents have overall justifiable reason to do.

The suggestion that rationality requires an appropriate response to beliefs about normative reasons corresponds to one analysis of subjective normative reasons, but subjective normative reasons have also been analysed as justified beliefs about what an agent has normative reason to do. Rationality might therefore be taken to require agents to respond appropriately to beliefs that they would be justified in forming about what they have normative reason to do.

9.4.2 Rationality requires an appropriate response to justified beliefs about overall reasons

One alternative to claiming that rationality involves responding appropriately to beliefs about reasons is to claim that rationality requires agents to respond appropriately to their justified beliefs about what they have reason to do. Cullity and Gaut suggest that rationality involves responding appropriately to subjective reasons. They write: 'an account of practical *rationality* must be given in terms of subjective reasons; one is practically rational to the extent to which one is guided by one's subjective normative reasons' (Cullity & Gaut, 1997, p. 2, italics in original). And, as shown in §6.2.1, Cullity and Gaut describe subjective reasons in a way that suggests that they are justified beliefs about normative reasons. Richard Joyce also discusses the claim that rationality involves appropriately responding to something like this form of subjective reason (2001, pp. 53-58). Parfit comes close to making this claim when he distinguishes between what an agent has overall normative reason to do and what it is rational for the agent to do. He claims that, 'the rationality of our desires and acts depends … on what we believe, or – given the evidence, ought rationally to believe' (Parfit, 1997, p. 99). Thus, Parfit can be read as claiming that rational agents act as they ought to believe they have reason to act, which comes close to claiming that agents are rational when they respond appropriately to justified beliefs about what they have overall reason to do. This analysis of rationality is more similar to my account of rationality than the analyses of rationality in terms of objective normative reasons or beliefs about reasons.

As follows from my discussion of subjective normative reasons in Chapter 6, I accept that if an agent responds appropriately to her justified beliefs about what she has normative reason to do, then she acts rationally. I accept that it is rational for agents to act appropriately in response to their justified beliefs about what they have reason to do whether those justified beliefs are about normative reasons of any kind or about objective normative reasons.[119] If an

---

[119] As argued in §4.4 and §6.3, saying that an agent has justifiable reason to act in some way is not the same as saying that an agent would be justified in believing that she has an objective normative reason to act in that way. An agent will act rationally if she acts appropriately in response to her justified beliefs about what she has

agent has a justified belief that she has overall normative reason to act in some way, she has justifiable reason to act in that way. So, I accept that if an agent responds appropriately to justified beliefs about what she has normative reason to do, she acts rationally.[120] I assume that the justification necessary for an agent to have a justified belief about what she has overall reason to do will amount to the same kind of justification necessary for an agent to have a justifiable reason for acting. This degree of justification will ensure that what is required for an agent to act rationally goes beyond mere consistency between means and ends.

John Broome would probably reject this claim that responding appropriately to justified beliefs about overall reasons is sufficient for acting rationally. According to Broome, various 'rational requirements' must be met for an action or belief to be rational. His examples of rational requirements include:

> *C1*. Necessarily, if you are rational, you do not believe p and believe not p.
>
> *C2*. Necessarily, if you are rational, you do not intend to F and intend not to F.
>
> *C3*. Necessarily, if you are rational then, if you believe p and you believe that if p then q, you believe q.
>
> *C4*. Necessarily, if you are rational then, if you intend to F and you believe you cannot F unless you G, you intend to G. (Broome, 2007a, p. 355)

Broome might argue that an agent could have a justified belief that she has overall reason to act some way, but not meet all the requirements of rationality, and if she failed to meet any of these requirements of rationality, he would argue that her action was not rational.

Broome admits that there may be exceptions to his rules of rationality, that is, situations where a rational agent fails to meet one of these rational requirements, but remains rational (2007a, pp. 361-362). For example, paraconsistent logicians may be rational and yet fail to satisfy *C1* (Broome, 2007a, p. 362).[121] Broome writes that he would be happy for *C1* to be weakened so that it can accommodate agents who are paraconsistent logicians; he revises it so that it says:

> [*C1R*.] Necessarily, if you are rational, you do not believe p and believe not p,
> unless you believe there are true contradictions. (Broome, 2007a, p. 362)

However, Broome points out that this weakening of *C1* could wrongly be taken as a sign that all his rational requirements could be similarly weakened, so that: 'we might think of

---

objective normative reason to do, but she can act rationally without having a justified belief that she has an objective normative reason to do anything. That is, an agent can act rationally even though she knows that her mental states are far from ideal. This may happen when an agent is aware that she does not know all the relevant facts about her situation or when she knows that her current set of beliefs and desires is incoherent.

[120] It doesn't follow that she ought to act in that way for all sorts of reasons. For example, if Nomy Arpaly's arguments about weakness of will are correct, the agent might justifiably believe that she has reason to act some way, but some other non-belief related part of her system might realise that she has more reason to act in some other way (2006).

[121] It is tempting to say that *C2* also needs to be revised for consequentialists who think that one should not aim to have one's actions maximise utility because that is not the best way to ensure that one's actions maximise utility. However, such consequentialists would claim that it was rational to 'intend to F and not intend to F'.

rationality as meeting your own standards, so that you are irrational only if you fail by your own standards' (2007a, p. 362). Broome argues that treating his weakening of *C1* to *C1R* in such a way would be a mistake. I agree.

The change from *C1* to *C1R* is plausible, but it is not made plausible because paraconsistent logicians only need to meet their own standards. Assume that Graham Priest and Richard Routley decided that arguing in support of true contradictions would improve their research rating, but that they would only be able to argue for this position if they believed there were true contradictions. They pay a hypnotist to instil a belief that there are true contradictions and then develop a system of paraconsistent logic. Priest and Routley now have *C1R* among their standards and not *C1*. But, in this situation Priest's and Routley's beliefs that there are true contradictions is irrational.[122] Paraconsistent logicians remain rational when they fail to meet *C1*, a requirement that the rest of us must meet, because they have good reasons for doubting that *C1* is always true. Priest and Routley did not ask to be hypnotised into accepting paraconsistent logic. They spent considerable time and thought developing the position and supporting it with argument (Priest, 2006; Priest & Routley, 1984; Routley, 1984). This explains why philosophers think that Priest and Routley are rational in spite of what many take to be a peculiar position on contradictions. The change from *C1* to *C1R* is only acceptable if the believer is justified in believing that there are true contradictions.

This argument is directly relevant to my claim that agents must respond appropriately to justifiable reasons to act rationally. John Broome claims that agents must abide by his rational requirements to be rational. If that is the case, then that is because Broome's rational requirements are statements of what we generally think people are justified in believing and intending. So, if he is right, this is at least consistent with my claim that agents must respond appropriately to justifiable reasons to act rationally. (I say 'at least consistent' because it is a leap to go from a set of rational requirements on internal states, to appropriate actions.) I accept that an agent's responding appropriately to her justified beliefs about what she has reason to do is sufficient for her to act rationally, but I do not accept that this is necessary for acting rationally. Basing rationality on 'justified beliefs' sounds as though an agent's believing she has reason to act in some way is key to her action being rational. What it is rational for an agent to do is not determined by her current beliefs, whether they are justified or unjustified. What matters is justification, not belief.

When an agent forms an intention to *F* because she has overall justifiable reason to form an intention to *F*, we assume that she has gone through a justifiable reasoning process. This reasoning process is part of what makes her action rational, but it is easy to forget that without this her action would not be rational. The word 'rational' is often used to refer to states of mind, and Broome sometimes uses it in this way in 'Does rationality consist in responding correctly to reasons?' (2007a). However, rationality involves getting all your ducks in a row, and while the ducks include mental states, they also include epistemic and practical reasoning, historic contexts, and states of the world. So, although the word 'rational'

---

[122] This claim assumes that beliefs aim at truth.

is used to refer to actions or cognitive attitudes, as in 'the agent acted rationally when she acted on her justified belief that she had reason to get her glasses', it does so based on the reasoning and responsiveness to circumstances that lies behind those actions or cognitive attitudes. What makes a belief a 'justified belief' is the way that the belief arose from reasoning, capabilities and circumstances. When I say that rationality involves responding appropriately to justifiable reasons, this connection between rationality and reasoning also gets hidden. But, what I mean when I give this analysis of rationality is that an agent is rational when she reasons appropriately about the aspects of her circumstances and nature that give her justifiable reason to act in certain ways. An agent is rational if, when she knows that she cannot remember which of two pill bottles contain her medication and which her partner's, she goes and gets her glasses to check the names on the bottles – or at least forms a sincere intention to do so. What makes her action rational is that acting in that way is an appropriate response to the various collections of *STATE, END, VAL*, and *SIG* that apply to her in her situation.

Agents can have reasons to act of which they are unaware. There are two ways that this can occur, through ignorance or through justifiable actions that are not responses to beliefs.

First, agents may be ignorant about what they have reason to do. This occurs when people have not thought about an issue yet and also when people ignore issues that they ought to have considered. We can talk about what it would be rational for an agent to do even when she has not reflected about her situation; we can talk about justifiable reasons an agent has for acting that she is totally ignorant of, but that it would be rational for her to act in response to. This requires only a minimal change to the 'justified beliefs about reasons' analysis of rationality, so that it says something like: 'an agent acts rationally if she forms those beliefs about what she has reason to do that she would be justified in forming, and acts appropriately in response to those beliefs'. The justification required would be determined by what was appropriate given her circumstances. However, there could be all sorts of beliefs about reasons that an agent might be justified in believing at an instant in time. It is reasonable to expect that time limitations would mean that engaging in the reasoning necessary to form one set of beliefs would prevent the agent forming another set of beliefs. So, it may be impossible to specify precisely what it would be rational for an agent to believe and do.

Second, as already discussed, an agent can act rationally without having a justified belief that she should act in that way. An agent could act rationally without that action being based on a belief; this is part of what Railton shows with his arguments about the significance of fluid agency and agent competence (2009).[123] An agent can have justifiable reason to act in a certain way, even though the reason and the action may be non-cognitive or cognitively inaccessible. Expert drivers can correct for skidding without being able to explain the manoeuvre they use to correct the car's direction, and musicians can spontaneously change key or alter timing to add interest. If we accept claims that know-how, that is, knowledge of how to act, is not reducible to knowledge that something is the case, an agent can act

---

[123] See §3.1 and §5.1.

rationally without acting appropriately in response to a justified belief about what she has reason to do.[124]

I have argued that rationality amounts to acting appropriately in response to justifiable reasons and that this requires more than an agent's acting on her justified beliefs about what she has reason to do. There is, however, another objection to the claim that rationality involves responding appropriately to reasons, an objection that applies to any account of reasons where the reason-explanans are not usually mental attitudes. Rationality is often said to be a relationship between attitudes, with practical rationality taken to be a relationship between beliefs or desires, and intentions. If, for example, James wants to inherit Uncle John's money and believes that drowning him is a means to get it, James acts rationally if he forms an intention to push Uncle John over the side. James's wanting, believing and intending determine what it is rational for him to do. In apparent contrast, normative reasons are taken to arise from a relationship between states of the world and attitudes. For example, Uncle John's wealth, last will and testament, and his being in the boat are reason-explanans for James to form an intention to push him overboard. James's reason-explanans are not said to be his beliefs about Uncle John's wealth, will, and Uncle John's being on the boat. If 'rationality' is taken to require mere consistency between an agent's ends and the means he thinks are available for achieving those ends, this analysis of rationality as a relationship between attitudes makes perfect sense; for means-end rationality, the factors that matter just are the agent's beliefs, desires and intentions. However, the arguments that aim to show that rationality is a relationship between mental attitudes also raise problems for my analysis of rationality in terms of justifiable reasons, because I claim that justifiable reasons are states of affairs, not mental attitudes.

9.5 Rationality as a relationship between attitudes

Ralph Wedgwood argues that rationality is about relationships between mental states, or 'attitudes', a position accepted by those who argue that rationality must involve an appropriate response by an agent to her own beliefs, or justified beliefs, about what she has overall reason to do (Broome, 2007a, p. 352; Kolodny, 2005, pp. 209-210; Scanlon, 2007; Wedgwood, 1999, 2002). Agents are said to be rational when they form intentions to act in agreement with their beliefs, not because they form intentions to act in agreement with the state of the world. Wedgwood writes:

> When we assess a choice or decision as rational or irrational, we are assessing it on the basis of its relation to the agent's beliefs, desires, and other such mental states—not on the basis of its relation to facts about the external world that could vary while those mental states remained unchanged. This also seems to be a special feature of rationality, in contrast to other ways of evaluating beliefs and decisions. All the other ways of evaluating beliefs and decisions—for example, as 'correct' or 'incorrect', 'advisable' or 'inadvisable', and so on—are externalist evaluations. What is distinctive of 'rationality' (at least as the term is most

---

[124] The claim that 'knowledge how' is not reducible to 'knowledge that' was made by Gilbert Ryle (Ryle & Dennett, 2000). Although it remains a contentious issue, Ryle's position is still well supported.

commonly used by philosophers) is that it is an internalist evaluation. (Wedgwood, 2002, p. 350)

James's belief that his uncle is wealthy makes it rational for him to push his uncle overboard, not James's uncle's wealth. Wedgwood uses arguments in support of internalism about epistemic justification to support his arguments about rationality. Internalists about epistemic justification, who are sometimes called 'epistemological internalists', hold that only features of agents determine whether beliefs are justified or unjustified.[125] According to Wedgwood, whenever agents correct or reaffirm beliefs, they do so in response to non-factive mental states or to relationships between mental states (2002).

Wedgwood uses an evil demon argument to support his claims. Compare the situations of two agents, Agnes and Ed, who have identical reasoning processes, 'experiences, intuitions, desires, and apparent memories' (Wedgwood, 1999, p. 118; See also Wedgwood, 2002, pp. 349-350). Agnes and Ed have both just experienced their doctor telling them to take a certain medicine. All the evidence available to Agnes and Ed suggests to them that this is a reason to take the medicine. It seems that they need to act in exactly the same way for their action to be rational. But, Ed's experiences were caused by an evil demon. Wedgwood concludes that because Agnes and Ed must act in the same way to act rationally, rationality is a relationship between mental states, for example, between Agnes's and Ed's experiences and their intentions.

Wedgwood's argument poses a problem for my account of justifiable reasons. Throughout this thesis I have argued that justifiable reason-explanans are states of affairs rather than mental states. Agnes and Ed seem to have the same justifiable reasons for acting; it seems appropriate for both Agnes and Ed to take themselves to have reason to take the medicine. Agnes's justifiable reason-explanans seems to be the state of affairs 'my doctor told me to take this medicine'. Ed's justifiable reason-explanans seems to be his evil-demon induced belief that his doctor told him to take the medicine. The simplest way to make Agnes's and Ed's justifiable reason-explanans the same kind of thing seems to be to say that justifiable reason-explanans are types of mental states – in this case, to say that Agnes's and Ed's justifiable reason-explanans are their experiences of their doctors telling them to take the medicine. But, if I alter my account of justifiable reasons by claiming that justifiable reason-explanans are types of mental states, it affects the relationship between justifiable reasons and objective normative reasons in a way that I would prefer to avoid. It seems undesirable for justifiable reason-explanans and objective normative reason-explanans to be metaphysically different kinds of things, that is, for justifiable reason-explanans to be types of mental states and objective normative reason-explanans to be states of affairs or propositions. Apart from this, there seems something right in the claim that it is usually states of the world that give agents normative reason for acting.

Wedgwood's argument is difficult to counter, but there are four possible responses. First, I could adopt the non-factive explanations move that Dancy makes for motivating reasons and

---

[125] William Alston outlines some forms of epistemological internalism in (Alston, 1986).

adapt it for justifiable reasons (2000, chs 6, 7 & 8). On a non-factive explanation of justifiable reasons, reasons are states of affairs, but those states of affairs need not exist. Ed's justifiable reason-explanans for taking the medicine would be the state of affairs 'my reliable doctor told me to take the medicine', but a third party might describe Ed's justifiable reason for acting by saying, 'as it seemed to Ed, his reliable doctor told him to take the medicine'. This differs from Dancy's use of non-factive explanation because justifiable reasons are normative and Dancy reserves this manoeuvre for motivating reasons, but this alone need not be a problem. However, as explained in §3.3, I am reluctant to incorporate non-factive explanation into my account of reasons because it seems to smuggle in mental states while claiming to eschew them. Second, I can argue that I am only concerned with so-called 'normal' circumstances. This is what I have done in the thesis up until now. This move requires me to concede that if my apparent interlocutors and I are deceived by an evil demon or in the Matrix, my account of reasons and rationality needs revising. Although this approach is common in philosophy, it involves side-stepping what some take to be an important issue. Third, I can rely on epistemological externalists to save my argument. I consider this approach briefly in §9.5.1. Finally, I can take the 'partners in crime' approach, and argue that if those who argue that justification is always internal are correct, everyone who argues that reasons are states of affairs or true propositions is affected to some degree. I would have to concede that justifiable reasons are cognitive attitudes. However, many of those who argue in support of objective normative reasons would also have to concede that objective normative reasons are cognitive attitudes. In §9.5.2, I argue for this position.

9.5.1 Epistemological externalism

Some epistemologists reject internalism about epistemic justification.[126] John Gibbons, for example, argues that justification is external, that is, he claims that it is not only factors internal to agents that justify beliefs (2006). Gibbons considers the situations of three omelette makers, call them 'Ojust' for 'Omelette Justified', 'Onote' for 'Omelette Note' and 'Ohist' for 'Omelette History'. All have the same internal states; they all recall buying eggs yesterday, believe that there are eggs in the fridge, and all three open the fridge door to get the eggs required for the omelette. However, while all three have the same internal states, their external circumstances differ. There is nothing in Ojust's history or environment to

---

[126] See also, Timothy Williamson's argument in *Knowledge and its Limits*, referred to by Wedgwood:

> I am not denying that knowledge ever plays a role in folk-psychological explanations. As Williamson has recently argued, knowledge does seem to play such a role in the explanation of certain actions [(Williamson, 2000, pp. 60-64, 75-88)]. For example, perhaps I keep on digging because I know that this mine contains gold. Believing, even truly believing, that it contains gold would not have been enough; for then I might have inferred this belief from a lemma whose falsity I might easily have discovered while digging, in which case I would have abandoned the belief and stopped digging. Here, however, the explanandum – my keeping on digging – consists in an agent's interacting with his environment in a certain way. It is only to be expected that the explanans – my knowing that the mine contains gold – will also consist in the agent's standing in a certain relation to his environment. This does not show that knowledge will figure in the explanation of an "internal" fact, such as the fact that a thinker comes to believe p at time t. An internal fact of this sort is surely more likely to have a correspondingly internal explanation. (Wedgwood, 2002, p. 362)

suggest there are no eggs in the fridge. There is a large note on Onote's fridge, a place where such notes are commonly put, saying 'I ate the eggs'. Ohist checked the fridge last night and noticed the eggs had been eaten, but is insufficiently reflective this morning to recall that information. Arguably only Ojust is justified in expecting there to be eggs in the fridge. Onote should have known that there were no eggs. Why? Because there was a note on the fridge door telling him that was the case. A difference between Ojust's and Onote's environment makes Ojust's intention to make an omelette justified when Onote's intention is unjustified. Ohist's intention to make an omelette is unjustified because of a combination of environment and history. Gibbons argues that it is the agents' external circumstances that explains the difference between what they are justified in doing.

I doubt that John Gibbons' thought experiment will convince an internalist to accept that justification is external. The internalist might respond that Onote is unjustified in intending to make an omelette because he failed to notice the mental representation of a note that the Matrix/evil demon ensured that Onote would become aware of if he switched his mental focus in a certain way. Similarly, the evil demon placed a simulated memory in Ohist of checking the fridge for eggs the night before, and Ohist should have searched his memory more carefully before the evil demon created the simulation of Ohist's morning kitchen experience. These are peculiar responses, but evil demon hypotheses are peculiar.

Gibbons' argument is only one of many arguments in support of externalism about epistemic justification, so whether justification is internal or external remains an open question (Pappas, 2005). Nevertheless, it seems sensible to try an alternative approach to avoiding the difficulty that justificatory internalism poses for my account of justifiable reasons and rationality.

9.5.2 What reasons does a brain in a vat have?

If justification is always internal, then justifiable reason-explanans are always attitudes. Strangely enough, this does not mean that justifiable reasons are never states of affairs, or if it does, objective normative reasons theorists must concede that objective normative reasons are often not states of affairs or true propositions. In epistemological internalists' arguments, agents deceived by evil demons, or living their lives as brains in vats, usually receive the same information about the state of what seems to them to be the world, their bodies and their minds, as the undeceived agents. As explained in §3.3, mental states, such as anger or depression, can be states of affairs that give agents reasons to act.

Some reasons are both states of affairs and mental states. It seems to Agnes and Ed that they touch a kettle and they feel their hand burning. The burning sensation is a state of affairs that gives them reason to stop touching the kettle. *VAL* and *SIG* associated with the burning sensation favour acting to stop the sensation. Each takes the hand away, the burning ceases, but it hurts. Agnes and Ed each experience putting the sore hand under running cold water and feeling the pain diminish. Each has a reason to form an intention to keep the hand under the cold water. The lessened pain is a state of affairs and a reason. In this example, Agnes and Ed have reason to form the same intentions to act, and the thing that gives them reason to act

is a state of affairs, namely, the pain and the lessening pain. The fact that Agnes is an agent in a world with real kettles and taps, while Ed is an agent whose life is a dream constructed by an evil demon, is irrelevant to what they have reason to form an intention to do. Agnes and Ed are rational if they, first, form an intention to move their hands, and, second, form intentions to keep them under the running water. In this example, I gave Agnes and Ed identical justifiable and objective normative reasons for intending to act. However, the example could be altered so that they have objective normative reasons to act, but not justifiable reasons to act, or vice versa. Assume that epistemological internalism is true. It would still be the case that some justifiable reasons and objective normative reasons could be states of affairs; rationality could sometimes involve acting appropriately in response to justifiable reasons that are states of affairs; rationality could require actions that go beyond any agent's current beliefs or desires; and what an agent has objective normative reason to do could differ from what it would be rational for the agent to do.

Situations where reasons are mental states are treated as the exception rather than the rule, but if the arguments of epistemological internalists succeed, justifiable reasons are always mental states.[127] I argue, however, that if epistemological internalists are right, justifiable reasons would always be mental states that served as states of affairs. Both Agnes and Ed have experiences of a stranger collapsing next to them. But in Agnes' case the stranger is a real person, while in Ed's case the experiences have been produced by an evil demon. The experience of a stranger's collapse gives Agnes and Ed justifiable reason to intend to act in the same way. As the situation is described, it is appropriate for both Agnes and Ed to take the experience of someone collapsing next to them to be a reason to move to help. I assume that Ed's history of experiences justifies his concluding that there is a stranger and the stranger needs help. So, it is appropriate to take Ed's evil-demon induced experience of a stranger collapsing to be a state of affairs that serves as a normative reason. In other words, even though we are referring to Ed's experiences rather than external states of affairs, those experiences can be considered from a perspective that is sufficiently idealised to make them normatively evaluable. If the evil demon makes Ed's experiences resemble normal human experiences, then not all of Ed's mental states will be normative reasons for acting. Like *ON*, Ed might sometimes fail to pay attention to the potential experiences the evil demon sends his way. When his experiences fail to correspond to the set of experiences that he would be justified in having, his mental states no longer serve as justifiable reasons for action.

It might seem at first that accounts of objective normative reasons can give a different response to this problem. An objective normative reasons theorist *might* claim that agents deceived by an evil demon have no reason to form an intention to help any stranger who seems to collapse next to them. Such a theorist would claim that a stranger's collapse in the real world is an objective normative reason for Agnes to try to help, but the simulated collapse in the illusory world is not an objective normative reason for Ed to try to help. Yet, someone who accepts an account of objective normative reasons is likely to accept that agents in the Matrix or deceived by evil demons have pro tanto normative reasons to form

---

[127] See, for example, (Dancy, 2000; Schroeder, 2008c).

intentions to act on a good portion of their experiences, and in those situations the agents' experiences serve as states of affairs.

Whether experiences inserted in agents' brains by evil demons provide objective normative reasons to help someone will depend on features of the reason relation other than *STATE* and *END*. Assume that the relationship between actions and effects in our world is replicated by evil demons' simulations. (This is not a peculiar assumption. The force of the epistemological internalist's argument stems from Agnes's and Ed's experiences being identical: according to the epistemological internalist, we could all be in Ed's situation.)

STATE	A state of affairs.	Agnes: The person collapses.
		Ed: Experience of a person collapsing.
END	Given the state of affairs, an action will yield an effect.	Agnes: The collapse is a sign that helping the person might improve his functioning.
		Ed: The experience of a person collapsing is a sign that helping the person might improve his functioning.
VAL	The effect has positive value.	Improving the functioning of any person who has collapsed in these circumstances would have some positive value.
SIG	The positive value of the action remains significant in spite of any negative values produced by the action.	Improving the functioning of a collapsed person in these circumstances would have positive value that would not conflict with other values to such a degree that it would be implausible to say there is reason to do it.

The contents of *VAL* and *SIG* determine whether Ed's experience of a stranger collapsing is a state of affairs that should serve as an objective normative reason. If actions are valuable purely because of their outcomes for people other than the agent, then Ed has no reason to form an intention to help the person he experiences collapsing. Torbjörn Tännsjö seems likely to support this response (2010).[128] This approach to reasons would make the objective normative reasons account of brain in a vat situations different from mine, but few objective normative reasons theorists discussed here would accept that the only value of actions is that realised by people other than the agent. In contrast, if experiences of acting can be valuable for agents themselves, whether because of character formation or positive affect, then the experience of helping someone who has collapsed could ground an objective normative reason relation even when the agent's experience of a stranger collapsing is the result of

---

[128] See Chapter 8 for exposition and discussion of Tännsjö's position.

being deceived by an evil demon. This would be the case even when the supporter of objective normative reasons was not concerned about the accessibility of normative reasons.

9.6 Conclusion

The account of justifiable reasons provided in this thesis is not just an account of normative reasons that real agents can use as the basis for their actions: justifiable reasons are also the reasons that explain what it means to be practically rational. I have argued that an agent acts rationally if and only if she acts appropriately in response to justifiable reasons. The hypothesis that rationality amounts to acting appropriately in response to normative reasons or beliefs about normative reasons is usually rejected. However, I showed that most of the objections considered here do not affect my claim about the relationship between rationality and justifiable reasons. The one argument that might affect my position is also likely to affect the theories developed by most objective normative reasons theorists.

The claim that agents act rationally whenever they act appropriately in response to normative reasons is commonly treated as a claim about the relationship between rationality and objective normative reasons. But, if an agent cannot know what she has objective normative reasons to do, it is possible for her to act rationally even though she is not acting as she has objective normative reason to act. Agents can always establish what they have justifiable reason to do, so this objection has no effect on my claim about the relationship between rationality and justifiable reasons.

An alternative to claiming that rationality amounts to acting appropriately in response to objective normative reasons is to claim that rationality amounts to agents acting appropriately in response to beliefs about what they have reason to do. I explained why the bootstrapping objection is usually taken to show that this analysis of rationality fails. However, the bootstrapping argument does not affect my claim that acting rationally amounts to acting appropriately in response to justifiable reasons. A second formulation of this proposed relationship between rationality and beliefs about reasons connects rationality to justified beliefs about reasons. This account bears more resemblance to my account of the reasons-rationality relation, and I almost agree with it. I agree that an agent acts rationally if she forms the beliefs about what she has reason to do that she would be justified in forming and acts appropriately in response to those beliefs. However, I argued that an agent can act rationally without acting appropriately in response to her justified beliefs about what she has reason to do. So explaining rationality in terms of justifiable reasons is more accurate than explaining rationality in terms of justifiable beliefs about reasons.

The claim that rationality is always a relationship between cognitive attitudes is more problematic for my claims. If rationality is always a relationship between cognitive attitudes, then it seems I must either give up my claim that justifiable reason-explanans are states of affairs and not usually psychological states, or give up my claim that rationality can be explained in terms of justifiable reasons. This claim that rationality is always a relationship

between cognitive attitudes relies on epistemological internalism. If epistemological internalists are correct, then my thesis needs to be rewritten to state that reason-explanans are usually experiences that it is appropriate to treat as states of affairs. But if epistemological internalists are correct, then anyone who gives an account of objective normative reasons where the agents' mental states contribute to the value of an action would also need to similarly rewrite their theory of reasons. And I take it that this would include many, if not all, of the objective normative reasons theorists discussed in this thesis.

10. Summing Up and Moving Forward

This thesis increases our understanding of reasons for action, and, in particular, our understanding of what it means for a real human agent – that is, a finite agent with limited information, time and capacity – to have reason to act in some way. I have developed an account of justifiable reasons – normative reasons for action that make sense of such commonsense truisms as:

> If choices are available, I can always act for a good reason.
>
> Something is not a reason for me to act if I cannot become aware of it.
>
> I am rational when I act appropriately in response to good reasons.
>
> Apart from trivial cases where my actions are unimportant, I deserve blame when I fail to determine what I have good reason to do.

Justifiable reasons are always accessible to real human agents. They are normative reasons that real human agents ought to take themselves to have pro tanto reason to act on.

I argued that normative reason-explanans – reasons that favour acting in some way – exist as reasons only because of their role within reason relations. And, I explained that although the word 'reason' can be used to refer to any of the relata within the reason relation, in this thesis I would take it to refer to states of affairs rather than the relationship between states of affairs, actions and outcomes, or the values that could be realised by acting in a certain way. An objective normative reason relation exists when a state of affairs (or proposition) indicates that acting in a certain way will have a certain outcome, and the outcome has significant positive value. The word 'value' is to be read very broadly, and having 'significant positive value' could be constituted by such things as taking pleasure in eating cake, feeling satisfaction after cleaning your desk or saying 'please' to act respectfully. And, the word 'outcome' should not be read as implying that only consequences matter; the relevant outcome could be something internal to the act itself, such as that the action shows respect for others. Different objective normative reasons theorists develop different accounts of what it means for a state of affairs, or proposition, to indicate that acting in a certain way will have a certain outcome and differing accounts of what it means for the outcome to have significant positive value. In contrast with objective normative reasons, an agent has a pro tanto justifiable reason to act in some way when it is appropriate for the agent to be aware of some state of affairs, and appropriate for her to treat that state of affairs as an indication that acting in a particular way will realise significant positive values. I do not attempt to classify justifiable reasons as 'subjective' or 'objective'. When researching normative reasons for this thesis I often found that claims about objectivity and subjectivity unhelpfully obscured the relevant issues.

My concept of justifiable reasons remains distinct from existing accounts of reasons without replacing existing coherent accounts of reasons. As I argued in Chapter 6, justifiable reasons cannot be reduced to, or explained in terms of, the accounts of subjective reasons discussed here. And, as argued in Chapter 7, although agents' motivational systems affect what it is appropriate for them to take to be good reasons for acting, my account of justifiable reasons cannot be reduced to, or explained in terms of, internal or external accounts of reasons, nor does it assume the truth of internalism or externalism about reasons. I showed that the same reason can be a justifiable, motivating or objective normative reason, but I also showed that justifiable reasons cannot be reduced to, or explained in terms of, motivating or objective normative reasons. Differences between objective normative reasons and justifiable reasons have been frequently pointed to throughout the thesis, and I have responded to potential objections that objective normative reasons theorists might make to justifiable reasons. However, I have not argued that the concept of objective normative reasons is incoherent, nor have I considered objections to any particular conception of objective normative reasons. Rather, I argued that justifiable reasons and objective normative reasons serve different functions within philosophy of action, functions which are complementary rather than conflicting.

This thesis adds at least one concept of reasons to the concepts already in use, and this may seem to multiply reasons beyond necessity. But my account of justifiable reasons is an attempt to give an account of a form of normative reasons that already plays key roles within everyday life and within philosophical arguments. I have shown that justifiable reasons are normative reasons that real agents have for acting and, I believe this is sufficient to make this concept of reasons worth serious consideration. But, as argued in Chapter 4, justifiable reasons also dovetail with correct accounts of rationality, self-regulation, and praise and blame. In Chapter 9, I argued that one important concept of rationality involves correctly responding to justifiable reasons. And, I showed that objections to attempts to explain rationality in terms of reasons fail to show that rationality cannot be explained in terms of justifiable reasons.

The research in this thesis has important implications for future practice. When philosophers develop arguments about normative reasons they should consider the intent of their arguments before deciding whether an account of justifiable reasons or objective normative reasons is best suited to the purpose at hand. Objective normative reasons may be useful for investigating objective values, by which I mean values that exist independently of any agent's ability to identify them as values. Anyone who, like Jonathan Dancy, aims to develop an account of objective normative reasons that agents can act on can use something like Dancy's epistemic filter to restrict the set of states of affairs that can be objective normative reasons to those that are within agents' 'epistemic kens', to borrow Peter Railton's term (Dancy, 2000, pp. 56-59, 65-66; 2004, pp. 158-159; Railton, 2008). Although, as I argued in Chapter 8, accounts of objective normative reasons can help us understand what real human agents have reason to do, the values that serve as reasons within justifiable reason relations will often not correspond to those that play the same role in objective normative reason relations. This means that anyone interested in investigating the values that are relevant to real human agents

should consider the relationship between justifiable reasons and values as well as conceptions of objective normative reasons. Objective normative reasons may also be useful for refining our understanding of what it means to be, in Michael Smith's terms, 'fully rational'. However, concluding that forms of rationality other than Smith's 'full rationality' are unrelated to reasons by considering only the relationship between rationality and motivating or objective normative reasons is too quick. As I argued in Chapter 9, other forms of rationality may be better explained in terms of justifiable reasons. Similarly, concluding that reasons for belief are fundamentally different from reasons for action because of differences between reasons for belief and objective normative reasons is too quick. The different roles that justification plays in theoretical and practical reason mean that a state of affairs can be a reason for believing p without being an objective normative reason to φ. This leads some objective normative reasons theorists to suggest that there is a fundamental difference in kind between reasons for action and reasons for belief. This disconnect between objective normative reasons and reasons for belief does not obtain for justifiable reasons and reasons for belief.

Perhaps surprisingly, my arguments about justifiable reasons also have implications for ethics. Justifiable reasons are, I argue, the only form of normative reason that agents can use as a basis for their actions. If an agent can act for an objective normative reason, that is only because it is also a justifiable reason for action. Moral reasons are a subset of normative reasons. So, the only moral reasons that agents can use as a practical basis for their actions are justifiable moral reasons. This relationship between justifiable reasons and ethics fits well with the relationship between justifiable reasons and praise and blame, self-regulation and rationality. To reiterate my earlier claims, objective normative reasons clearly matter to arguments about ethics even though justifiable moral reasons are the only moral reasons on which agents can base their actions. That they do matter is evident in the relationship between objective normative reasons and value. But when philosophers want to discuss moral reasons that real human agents have for acting, they should focus on justifiable reasons, not objective normative reasons.

My claim that real human agents' moral reasons are justifiable reasons leaves me vulnerable to an objection also faced by internalists: it will sometimes be inappropriate for real human agents to take themselves to have justifiable reason to act in ways that many other people take to be morally right, or, as importantly, at times it will be appropriate for real human agents to take themselves to have justifiable reason to act in ways that many other people are certain are morally wrong (Dancy, 2000, pp. 16-17; Finlay, 2006, pp. 4, 6-7; Scanlon, 1998, pp. 363-373). Just as people in the Middle Ages had no justifiable reason to object to piling up dead bodies next to food markets, there are probably people alive now who have no justifiable reason to buy eggs from cage-free hens, or to cut the arms off albinos to make what they take to be medicines. I bite the bullet and accept this consequence, but, as I argue in §7.2.4, I do not accept that this makes justifiable reasons – or ethics – worryingly relativistic.

Although it may not be obvious, the fact that justifiable reasons are the only normative reasons that real human agents can use as the basis for their actions points to the importance of community. Human limitations delineate what agents have justifiable reason to do. And, as

the research on empathy and motivation discussed in this thesis shows, we do not always know our own limitations. However, human agents exist within communities, and these communities significantly increase the resources available for learning what we have reason to do. This makes it particularly important that, as members of communities, we try to stay aware of the ways in which we can deliberate with other people about our reasons for acting. Individualistic accounts of autonomy – that is, accounts of autonomy that emphasise individual reasoning and decision-making at the cost of collective reasoning and decision-making – risk limiting what agents have justifiable reason to do in ways that are undesirable and unnecessary. Those who argue that we can have autonomy without individualism do us a service by encouraging people to see the importance of community to autonomy, and hence of community to agency.[129]

Many possibilities for future projects arise from the work in this thesis. First, I argued that what agents have justifiable reason to do is limited by what it is epistemically and practically possible and appropriate for agents to learn about the circumstances. A more developed conception of what it means for it to be 'practically possible and appropriate' for an agent to learn something about her circumstances would not significantly alter my concept of justifiable reasons, but might usefully refine my arguments. I am particularly interested in considering the role that so-called 'know-how', that is, procedural knowledge as opposed to propositional knowledge, plays in determining what agents have normative reason to do, however emotions and moods also affect what it is possible and appropriate to expect of agents, so it is worth considering their effect on agents' justifiable reasons. Second, the parallels between reasons for belief and justifiable reasons suggest that arguments about reasons for belief could be used to refine and develop the account of justifiable reasons given in this thesis, and arguments about justifiable reasons may be similarly useful for refining accounts of reasons for belief. This would be a big project, but seems worth undertaking. Third, I have not considered the nexus between justifiable reasons and work on decision theory, nor have I investigated whether work on decision theory could help refine my work on justifiable reasons. The distinction between justifiable reasons and objective normative reasons is not a distinction between a decision procedure and an account of truth-making factors. My account of justifiable reasons is an account of the truth conditions for an agent having justifiable reason for acting, not a decision procedure. So, work on decision theory is orthogonal to my work in this thesis. Nevertheless, the analogies between the aims of decision theorists and my aims suggest further research into the relationship would be worthwhile.

---

[129] See, for example, G. M. Stirrat and R. Gill, 'Autonomy in Medical Ethics after O'Neill', and the references to other works in this article (2005).

References

Adams, E. M. (1993). Rationality and Morality. *Review of Metaphysics, 46*(4), 683-697.

Alston, W. P. (1986). Internalism and Externalism in Epistemology. *Philosophical Topics, 14*(1), 179-221.

Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social Influences on Creativity: The Effects of Contracted-For Reward. *Journal of Personality and Social Psychology 50*(1), 14-23.

Arpaly, N. (2006). *Merit, Meaning, and Human Bondage: An Essay on Free Will*: Princeton University Press.

Audi, R. (2001). *The Architecture of Reason: The Structure and Substance of Rationality*. Oxford: Oxford University Press.

Benner, P. (2001). *From Novice to Expert: Excellence and Power in Clinical Nursing Practice*: Prentice Hall.

Bratman, M. (1987). *Intention, Plans, and Practical Reason*. Cambridge, Massachusetts: Harvard University Press.

Brierre de Boismont, A. J. F. (1859). *On Hallucinations: A History and Explanation of Apparitions, Visions, Dreams, Ecstasy, Magnetism, and Somnambulism* (R. T. Hulme, Trans.). London: Savill and Edwards.

Broome, J. (1999). Normative Requirements. *Ratio, 12*(4), 398-419.

Broome, J. (2001). Are Intentions Reasons? And How Should We Cope with Incommensurable Values? In C. W. Morris & A. Ripstein (Eds.), *Practical Rationality and Preference: Essays for David Gauthier* (pp. 98-120). Cambridge: Cambridge University Press.

Broome, J. (2006). Reasons. In R. J. Wallace, P. Pettit, S. Scheffler & M. Smith (Eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (pp. 28-55). Oxford: Oxford University Press.

Broome, J. (2007a). Does Rationality Consist in Responding Correctly to Reasons? *Journal of Moral Philosophy, 4*(3), 349-374.

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

Chang, R. (2006). Can Desires Provide Reasons for Action? In R. J. Wallace, P. Pettit, S. Scheffler & M. Smith (Eds.), *Reason and Value:Themes from the Moral Philosophy of Joseph Raz*. New York: Oxford University Press.

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

Cullity, G., & Gaut, B. N. (1997). Introduction. In G. Cullity & B. N. Gaut (Eds.), *Ethics and Practical Reason* (pp. 1-27). Oxford: Oxford University Press.

Dancy, J. (2000). *Practical Reality*. Oxford: Oxford University Press.

Dancy, J. (2003). Replies. *Philosophy and Phenomenological Research, 67*(2), 468-490.

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

Dancy, J. (2009). Reasons and Rationality. In S. Robertson (Ed.), *Spheres of Reason: New Essays in the Philosophy of Normativity* Oxford: Oxford University Press.

Darwall, S. (1997). Reasons, Motives, and the Demands of Morality: An Introduction. In S. Darwall, A. Gibbard & P. Railton (Eds.), *Moral Discourse and Practice: Some Philosophical Approaches*. Oxford: Oxford University Press.

Darwall, S. (2006a). Morality and Practical Reason; A Kantian Approach. In D. Copp (Ed.), *The Oxford Handbook of Ethical Theory* (pp. 282-320). Oxford: Oxford University Press.

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

Davidson, D. (1963). Actions, Reasons and Causes. *The Journal of Philosophy, 60*(23), 685-700.

Eccles, J. S., & Wigfield, A. (2002). Motivational Beliefs, Values, and Goals. *Annual Review of Psychology, 53*(1), 109-132.

Eva, K. W. (2010). On the Relationship Between Problem-Solving Skills and Professional Practice. In C. Kanes (Ed.), *Elaborating Professionalism: Studies in Practice and Theory* (pp. 17-34). Dordrecht: Springer.

Fantl, J., & McGrath, M. (2009). *Knowledge in an Uncertain World*. Oxford: Oxford University Press.

Fine, C. (2010). *Delusions of Gender*. London: Icon Books.

Finlay, S. (2006). The Reasons that Matter. *Australasian Journal of Philosophy, 84*(1), 1-20.

Gert, J. (2008). Putting Particularism in its Place. *Pacific Philosophical Quarterly, 89*(3), 312-324.

Gibbons, J. (2006). Access Externalism. *Mind, 115*(457), 19-39.

Glucksberg, S. (1962). The Influence of Strength of Drive on Functional Fixedness and Perceptual Recognition. *Journal of Experimental Psychology, 63*(1), 36-41.

Goldman, A. H. (2009). *Reasons from Within: Desires and Values*. Oxford: Oxford University Press.

Greenspan, P. (2007). Practical Reasons and Moral 'Ought' In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (pp. 172-194). Oxford: Oxford University Press.

Hare, R. M. (1972). Wanting: Some Pitfalls. In R. M. Hare (Ed.), *Practical Inferences* (pp. 44-58). Berkeley and Los Angeles, CA: University of California Press.

Hawthorne, J., & Stanley, J. (2008). Knowledge and Action. *Journal of Philosophy, 105*(10), 571-590.

Ickes, W. (2003). *Everday Mind Reading: Understanding What Other People Think and Feel*. Amherst, NY: Prometheus Books.

Ickes, W., Gesn, P., & Graham, T. (2000). Gender Differences in Empathic Accuracy: Differential Ability or Differential Motivation? *Personal Relationships, 7*, 95-109.

Johnson, L., & Johnson, N. (1987). Knowledge Elicitation Involving Teachback Interviewing. In A. L. Kidd (Ed.), *Knowledge Acquisition for Expert Systems; An Introductory Framework*. New York: Plenum Press.

Jones, K. (2003). Emotion, Weakness of Will, and the Normative Conception of Agency. In A. Hatzimoysis (Ed.), *Philosophy and the Emotions* (pp. 181–200). Cambridge: Cambridge University Press.

Joyce, R. (2001). *The Myth of Morality*. Cambridge: Cambridge University Press.

Kagan, S. (1989). *The Limits of Morality*. Oxford: Oxford University Press.

Kant, I., & Ellington, J. W. (1993). *Grounding for the Metaphysics of Morals ; With, On a Supposed Right to Lie Because of Philanthropic Concerns*: Hackett Publishing Company.

Klein, K. J. K., & Hodges, S. D. (2001). Gender Differences, Motivation and Empathic Accuracy: When it Pays to Understand. *Personality and Social Psychology Bulletin, 27*(6), 720-730.

Kolodny, N. (2005). Why Be Rational? *Mind, 114*(455), 509-563.

Korsgaard, C. (1986). Scepticism about Practical Reason. *The Journal of Philosophy, 83*(1), 5-25.

Korsgaard, C. (1996). *Creating the Kingdom of Ends*: Cambridge University Press.

Lemos, R. M. (1995). *The Nature of Value: Axiological Investigations*. Gainesville, FL: University Press of Florida.

Lillehammer, H. (2010). Facts, Ends, and Normative Reasons. *Journal of Ethics, 14*(1), 17-26.

Lord, E. (2010). Having Reasons and the Factoring Account. *Philosophical Studies, 149*(3), 283-296.

Loux, M. J. (2002). Propositions and Their Neighbours. In M. J. Loux (Ed.), *Metaphysics: A Contemporary Introduction* (2nd, revised ed., pp. 139-173). Abingdon, Oxon: Routledge.

Macklin, R. (1968). Norm and Law in the Theory of Action. *Inquiry: An Interdisciplinary Journal of Philosophy, 11*(1), 400-409.

Mason, C. (2006). Internal Reasons and Practical Limits on Rational Deliberation. *Philosophical Explorations, 9*(2), 163-177.

McDowell, J. (1979). Virtue and Reason. *The Monist, 62*(3), 331-350.

McDowell, J. (1995). Might there Be External Reasons. In J. E. J. Altham & R. Harrison (Eds.), *World, Mind, and Ethics* (pp. 68-85). Cambridge: Cambridge University Press.

McDowell, J. (2006). Response to Philip Pettit and Michael Smith. In C. Macdonald & G. Macdonald (Eds.), *McDowell and His Critics*. Malden, MA, USA: Blackwell Publishing Ltd.

Mele, A. (2007). Reasonology and False Beliefs. *Philosophical Papers, 36*(1), 91-118.

Northcutt, W. (2007). *The Darwin Awards Omnibus*. Camberwell, Victoria: Penguin.

Pappas, G. (2005). Internalist vs. Externalist Conceptions of Epistemic Justification. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Stanford: Stanford University

Parfit, D. (1997). Reasons and Motivation. *Proceedings of the Aristotelian Society, 71*, 99-130.

Parfit, D. (2001). Rationality and Reasons. In J. J. Dan Egonsson, Björn Petterson, Toni Rønnow-Rasmussen (Ed.), *Exploring Practical Philosophy: From Action to Values* (pp. 17-39). Burlington, VT: Ashgate.

Potrč, M., & Strahovnik, V. (2004). *Practical Contexts*. Heusenstamm: Ontos Verlag.

Priest, G. (2006). *Doubt Truth to Be a Liar*. Oxford: Oxford University Press.

Priest, G., & Routley, R. (1984). Introduction: Paraconsistent Logics. *Studia Logica: An International Journal for Symbolic Logic, 43*(1-2), 3-16.

Railton, P. (2008). Reply to John Skorupski. *Utilitas, 20*, 230-242.

Railton, P. (2009). Practical Competence and Fluent Agency. In D. Sobel & S. Wall (Eds.), *Reasons for Action* (pp. 81-115). Cambridge: Cambridge University Press.

Raz, J. (1999). Explaining Normativity: On Rationality and the Justification of Reason. *Ratio, 12*(4), 354-379.

Rorty, A. (2009). On Being Rational. *Ratio, 22*(3), 350-358.

Ross, D. (2003). The Right and the Good. In P. Stratton-Lake (Ed.). New York: Oxford University Press.

Routley, R. (1984). The American Plan Completed: Alternative Classical-Style Semantics, without Stars, for Relevant and Paraconsistent Logics. *Studia Logica: An International Journal for Symbolic Logic, 43*(1-2), 131-158.

Ryle, G., & Dennett, D. C. (2000). *The Concept of Mind*: University of Chicago Press.

Sayre-McCord, G. (1997). The Metaethical Problem. *Ethics, 108*(1), 55-83.

Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Scanlon, T. M. (2007). Structural Irrationality. In G. Brennan, R. Goodin, F. Jackson & M. Smith (Eds.), *Common Minds: Essays in Honor of Philip Pettit* (pp. 84-103). Oxford: Oxford University Press.

Schroeder, M. (2007). The Humean Theory of Reasons. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (Vol. 2, pp. 195-219). Oxford: Oxford University Press.

Schroeder, M. (2008a). *Being For: Evaluating the Semantic Program of Expressivism*: Oxford University Press.

Schroeder, M. (2008b). Having Reasons. *Philosophical Studies, 139*(1), 57-71.

Schroeder, M. (2008c). *Slaves of the Passions*. Oxford: Oxford University Press.

Schroeder, M. (2009). Means-End Coherence, Stringency, and Subjective Reasons *Philosophical Studies, 143*(2), 223-248.

Schroeder, M. (2011). What Does it Take to 'Have' a Reason? In A. Reisner & A. Steglich-Petersen (Eds.), *Reasons for Belief*: Cambridge University Press.

Setiya, K. (2004). Against Internalism. *Nous, 38*(2), 266-298.

Setiya, K. (2007). *Reasons Without Rationalism*. Princeton: Princeton University Press.

Shaw, W. H. (1995). *Moore on Right and Wrong: The Normative Ethics of G.E. Moore*. Dordrecht: Kluwer.

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting Adolescent Cognitive and Self-Regulatory Competencies from Preschool Delay of Gratification: Identifying Diagnostic Conditions. *Developmental Psychology, 26*(6), 978-986.

Sie, M., Slors, M., & Brink, B. v. d. (2004). Introduction: The Problem of Reasons of One's Own. In M. Slors, M. Sie & B. v. d. Brink (Eds.), *Reasons of One's Own*. Aldershot: Ashgate.

Skorupski, J. (1997). Reasons and Reason. In B. G. Garrett Cullity (Ed.), *Ethics and Practical Reason*. Oxford: Oxford University Press.

Skorupski, J. (2006). Propositions about Reasons. *European Journal of Philosophy, 14*(1), 26-48.

Skorupski, J. (2009). The Unity and Diversity of Reasons. In S. Robertson (Ed.), *Spheres of Reason: New Essays in the Philosophy of Normativity*. Oxford: Oxford University Press.

Smith, M. (1989). Dispositional Theories of Value. *Proceedings of the Aristotelian Society, Supp Vol. 63*, 89-111.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Smith, M. (1995a). Internal Reasons. *Philosophy and Phenomenological Research, 55*, 109-131.

Smith, M. (1995b). Internalism's Wheel. *Ratio (New Series) 8*(3), 277-302.

Smith, M. (1996). Normative Reasons and Full Rationality: Reply to Swanton. *Analysis, 56*(3), 9.

Smith, M. (1997a). In Defense of "The Moral Problem": A Reply to Brink, Copp, and Sayre-McCord. *Ethics, 108*(1), 84-119.

Smith, M. (1997b). A Theory of Freedom and Responsibility. In B. G. Garrett Cullity (Ed.), *Ethics and Practical Reason*. Oxford: Oxford University Press.

Smith, M. (2003). Rational Capacities. In C. Tappolet & S. Stroud (Eds.), *Weakness of Will and Varieties of Practical Irrationality*. Oxford: Oxford University Press.

Smith, M. (2004). Philosophy and Commonsense: The Case of Weakness of Will. In M. Smith (Ed.), *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics* (pp. 56-72). Cambridge: Cambridge University Press.

Sobel, D. (2001). Subjective Accounts of Reasons for Action. *Ethics, 111*(3), 461-492.

Stirrat, G. M., & Gill, R. (2005). Autonomy in Medical Ethics after O'Neill. *Journal of Medical Ethics, 31*(3), 127-130.

Swanton, C. (1996). Is the Moral Problem Solved? *Analysis, 56*(3), 155–160.

Tännsjö, T. (2010). *From Reasons to Norms: On the Basic Question in Ethics*. Dordrecht, Netherlands: Springer

Wallace, R. J. (2003a). Review of Richard Joyce, *The Myth of Morality*. *Notre Dame Philosophical Reviews,* (11 Apr). Retrieved from http://ndpr.nd.edu/review.cfm?id=1301

Wallace, R. J. (2003b). Review: Explanation, Deliberation, and Reasons. *Philosophy and Phenomenological Research, 67*(2), 429-435.

Wallace, R. J. (2006). *Normativity and the Will: Selected Essays on Moral Psychology and Practical Reason*: Clarendon Press.

Wallace, R. J. (2006). Reason, Desire, and the Will. In R. J. Wallace (Ed.), *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason* (pp. 63-70). Oxford: Oxford University Press.

Wallace, R. J. (2007). Moral Motivation. In J. Dreier (Ed.), *Contemporary Debates in Moral Theory* (pp. 182-196). Malden MA: Blackwell.

Wechsler, D. (2002). *Wais-III, Wechsler Adult Intelligence Scale, Third Edition*: Psychological Corporation.

Wedgwood, R. (1999). The A Priori Rules of Rationality. *Philosophy and Phenomenological Research, 59*(1), 113-131.

Wedgwood, R. (2002). Internalism Explained. *Philosophy and Phenomenological Research, 65*(2), 349-369.

Wedgwood, R. (2008). Reasons Without Rationalism, by Kieran Setiya. *Mind, 117*(468), 1130-1135.

Wentzel, K., & Wigfield, A. (1998). Academic and Social Motivational Influences on Students' Academic Performance. *Educational Psychology Review, 10*(2), 155-175.

Williams, B. (1981a). Internal and External Reasons *Moral Luck* (pp. 101-113). Cambridge: Cambridge University Press.

Williams, B. (1981b). Persons, Character and Morality. In B. Williams (Ed.), *Moral Luck: Philosophical Papers 1973-1980* (pp. 1-19). Cambridge: Cambridge University Press.

Williams, B. (1995a). Internal Reasons and the Obscurity of Blame *Making Sense of Humanity*. Cambridge: Cambridge University Press.

Williams, B. (1995b). Replies. In J. E. J. Altham & R. Harrison (Eds.), *World, Mind, and Ethics*. Cambridge: Cambridge University Press.

Williams, B. (2001). Postscript: Some Further Notes on Internal and External Reasons. In E. Millgram (Ed.), *Varieties of Practical Reasoning*. Cambridge, Massachusetts: MIT Press.

Williamson, T. (2000). *Knowledge and its Limits*: Oxford University Press.

Wright, G., & Ayton, P. (1987). Eliciting and Modelling Expert Knowledge *Decision Support Systems, 3*(1), 13-26.