

**Making Moral Judgements**  
**Internalism and Moral Motivation**

**Matthew James Denver Millar**

**Doctor of Philosophy**

**The University of Edinburgh**

**2005**

# Contents

---

<b>Abstract</b>	iv
<b>Declaration</b>	v
<b>Acknowledgements</b>	vi
<b>Introduction</b>	1
<b>1. Internalism and the Problem of Moral Motivation</b>	6
1. Introduction	6
2. Non-cognitivist Internalism	13
3. Content-based Arguments for Internalism	17
4. Internalism and Justifying Reasons	27
5. Action-guiding Content and Motivating States	34
6. Accommodating the Internalist Intuition	37
7. Conclusion	43
<b>2. Externalism and Moral Motivation</b>	44
1. Introduction	44
2. Externalist Models of Moral Motivation	47
3. The Externalist Argument from Amoralism	50
4. Smith's Anti-Externalism	55
5. Externalism and the Two Views of Morality	77
6. Conclusion	84
<b>3. Speech Acts and Moral Motivation</b>	86
1. Introduction	86
2. Moral Motivation and Moral Language	88
3. The Structure of Speech Acts: Illocutionary Force and Propositional Content	90
4. Assertion and Compound Speech Acts	94
5. The Nature of Illocutionary Acts	97
6. How to Make a Moral Judgement	100
7. Copp's Discourse Internalism	105



8. Evaluative Predicates and Force Indication	113
9. Review of the Argument So Far and Potential Difficulties for the Speech Act Approach	119
10. Conclusion	124
<b>4. Normative Expectation and the Participant Point of View</b>	125
1. Introduction	125
2. Failures of Motivation and Sincerity Conditions	125
3. Normative Expectation and Universalizability	131
4. Unwarranted Sanctioning Emotions and Responses	133
5. Holding Someone to an Expectation and the Reactive Emotions	140
6. Susceptibility to the Reactive Emotions and the Mutual Dependence Claim	145
7. Moral Judgement, Normative Competence and the Amoralist	152
8. Bare Moral Belief	162
9. Conclusion	166
<b>5. Moral Thought, Intentional Action and the Desire/Belief Thesis</b>	168
1. Introduction	168
2. Moral Thought and Moral Language	168
3. Conative Motivation and Normative Motivation	174
4. Normative Motivation, Intentional Action, and the Desire/Belief Thesis	180
5. Internalism, Practical Force, and the Role of Conative States	189
6. Normative Content and Internal Reasons	194
7. General Internalist Disquiet	202
8. Internalism about Moral Requirements and Motivation	206
9. Conclusion	209
<b>Conclusion</b>	211
<b>Bibliography</b>	213

## Abstract

---

*Making Moral Judgements* is about the relation between moral judgements and motivation. It addresses an apparent tension between the internalist view that there is a necessary connection between moral judgements and motivation and the cognitivist view that such judgements express beliefs about how the world is morally. This thesis argues that to resolve this tension we need first to distinguish the act of making a moral judgement from the content of a moral judgement. Act internalism asserts a necessary connection between the act of making a moral judgement and motivation. This is perfectly consistent with there being no connection between the truth of moral propositions – the content of moral judgements – and motivation. The act internalist approach is developed using speech act theory. Speech acts that do more than simply state facts or express motivating states are ubiquitous in our linguistic practice. Moral judgements can be construed as a type of compound speech act that involves assertion and motivation. This approach, it is argued, can help us better understand the complexities of moral motivation and of moral practice. On the speech act approach, in making a moral judgement an agent goes beyond description or cognition in holding herself and others to account with regard to a moral requirement. To be able to do this, the agent must be generally susceptible to a range of reactive attitudes that make up the point of view of normative participation. It is in relation to this participant point of view that we can account for the capacity of agents to be motivated by moral considerations. And it is with regard to this point of view that internalism is best understood as an expression of our interested or participatory relation to moral deliberation and moral practice.

## Declaration

---

I certify that this thesis does not incorporate any material which has been submitted for the award of any other degree or professional qualification in any university; and that to the best of my knowledge and belief it does not contain any material previously written or published by another person except where due reference is made in the text.

Department of Philosophy  
The University of Edinburgh  
January 2006

## Acknowledgements

---

Firstly I express my sincere thanks to my supervisors Rae Langton, Richard Holton and Mike Ridge. I am also indebted to Rodney Allen for many helpful comments.

Many thanks to participants in the University of Edinburgh Department of Philosophy post-graduate seminars.

Financial support for this research was provided by a Faculty of Arts Group Research Scholarship, an Overseas Research Students Award, and two J Warren Macalpine Scholarships.

Family and friends who have provided encouragement and support include Jean-Louis Hudry, Judith and Warwick Millar, Samantha Millar, and Sarah and David Pearce.

In memory of Doris Davies.

Special thanks to Wendy Rogers for her love and support across two hemispheres.

## Introduction

---

The aim of this thesis is to understand the nature of the relation between moral judgements and motivation in order to defend the view that moral judgements are both motivating and fact-stating. My point of departure is the non-cognitivist anti-cognitivist argument. This says that if moral judgements are of conceptual necessity connected to motivation, then the primary function of such judgements cannot be to state facts or express beliefs. The view that there is a connection of conceptual necessity between moral judgements and motivation is known as moral judgement internalism. In chapter one, I claim that the anti-cognitivist argument is invalid and that there is no quick way of dismissing the union of internalism and cognitivism. Nonetheless, I stress that we should take seriously the non-cognitivist view that moral judgements in some sense go beyond mere belief or description. I suggest we think of internalism as an expression of our interested or participatory relation to moral deliberation and practice.

According to the non-cognitivist version of internalism, the connection between moral judgements and motivation holds at the level of linguistic function rather than at the level of truth-conditional content. This suggests a distinction between two senses of 'moral judgement'. Sometimes when we talk about moral judgements our focus is on what is judged – the moral proposition. For example, we might say that her judgement that stealing is wrong is true. In such cases, 'moral judgement' refers to the content of the judgement. But 'moral judgement' can also refer to the act of making a judgement. We might say, for example, that her judgement about stealing was vehement or ill-timed. In such cases it seems that our focus is on the act of judging rather than on the proposition judged. I argue that recognising this distinction is crucial to understanding the debate between moral judgement internalists and externalists ( those who deny a necessary connection between moral judgements and motivation).

Those who defend the combination of cognitivism and internalism tend to assume that the internal connection to motivation must be secured via moral belief. However, the notion of a necessary connection between belief and motivation is difficult to sustain in light of the orthodox view of belief as motivationally inert. This view forms part of a theory of motivation and action explanation called the desire/belief thesis. It says that beliefs can only motivate with the help of essentially motivating states such as desire. Cognitivists sometimes attempt to accommodate internalism and the desire/belief thesis by including motivation in the content of moral judgements. I reject 'content internalism' as an interpretation of moral judgement internalism because it fails to secure a necessary connection between moral judgements and motivation. It is also inconsistent with the widely accepted doctrine that moral requirements apply to agents regardless of their contingent and particular desires. I claim that to defend cognitivist internalism we need to develop the view that there is a necessary connection between the *act* of making a moral judgement and motivation.

In chapter two, I explore the possibility that internalism is misguided and that a plausible version of moral cognitivism is better served by adopting the thesis of moral judgement externalism. Externalists reject the view that there is a necessary connection between genuine moral judgements and motivation. They argue that moral judgements motivate, when they do, because of the presence of an independent and contingent desire to do what is right or, perhaps, because of a desire the agent happens to have to do the act judged right. Part of the reason externalists view internalism as a threat to cognitivism and, by extension, to realism in ethics is that they accept the non-cognitivist anti-cognitivist argument. I claim that externalism is far less compelling once this argument is rejected. I also attempt to show how the failure to distinguish clearly between the act and content of moral judgements obscures the nature of the debate between moral judgement internalists and externalists. I conclude that externalist explanations of the action-guiding and motivating features of moral judgements are unconvincing, given their claim that the motivating force and justifying authority of such judgements are only contingently related to the nature of moral requirements. I also raise some doubts about the

standard view of the role of desires and beliefs in motivation and the explanation of action.

In chapter three, I use Searle's theory of speech acts to develop the view that moral judgement internalism is a doctrine about judgement acts. 'Act internalism' allows us to avoid locating the connection between moral judgements and motivation at the level of the content of moral judgements or at the level of an external and independent desire to do the right thing. The speech act approach is one way of interpreting act internalism. Speech acts that involve both cognitive and non-cognitive states are common in our linguistic practice. I argue that construing moral judgements as a type of compound speech act allows us to account for the fact-stating and the motivating aspects of such judgements.

A potential problem with the speech act model is that it seems unable to account for failures of moral motivation that result from depression, accidie, and other kinds of mood disorders. If moral judgements entail motivation, then those who fail to be motivated because of such disorders are unable to make genuine moral judgements. However, agents who fail to be motivated by what they judge to be right are often well aware of their failings. In at least some of these cases the agents appear to be capable of making genuine moral judgements despite their lack of motivation. In chapter four, I attempt to refine my speech act account of moral judgements to incorporate the problem of failures of motivation. I argue that an agent can be sincere in her moral judgement even if she fails to be motivated by it, so long as she holds herself and others to account with regard to the requirement expressed by the judgement.

The notion of holding to account or 'normative expectation' I adapt from the work of Peter Strawson and R.J. Wallace. Specifically, I follow Strawson and Wallace in attempting to understand the stance of holding someone to account in relation to a range of sanctioning or 'reactive' attitudes. I develop my own view of the relation between normative expectation and the reactive attitudes. I argue that to be capable of holding someone to account an agent must be generally susceptible to the reactive

attitudes and sanctioning responses that characterise the point of view of normative engagement. This allows me to assert a necessary connection between making a moral judgement and being susceptible to a range of motivating attitudes, while accommodating the view that moral judgements function to express beliefs and state facts. In this way my approach secures both moral judgement internalism and moral cognitivism.

In chapter five I address the criticism that, while the speech act approach provides a version of internalism about moral language, it does not address the deeper question of whether *thinking* something morally right entails motivation. The speech act account needs to explain in what sense the rules governing moral speech acts have implications for moral psychology. I claim that this point misconstrues the nature of the connection between moral thought and moral language. I concede that it is logically possible to have moral beliefs without motivation; however, I argue that because of the interdependence between moral practice and moral language, the rules governing moral speech acts play a crucial role in the nature of moral deliberation and moral thought.

Next, I turn my attention to the account of moral motivation suggested by the speech act analysis of moral judgements. I argue that the picture of moral motivation that emerges is far too complex to be accommodated by the standard desire/belief model of motivation. Motivation seems to figure in the reactive attitudes entailed by agents' moral judgements as well as, in some sense, in the agents' responses to the justifying content of their judgements. I argue that we need to distinguish between what I term 'normative motivation' and 'conative motivation'. This is basically a distinction between motivation by justifying considerations and motivation by conative states. The distinction is used to address the issue of whether moral beliefs are capable of motivating and the status of the desire/belief model of motivation and action explanation. I argue that moral beliefs are ideally suited to convey the considerations that normatively motivate agents to act and that desires are ideally suited to their role as states of being motivated.



I then claim that any plausible account of intentional action must accommodate the distinction between normative and conative motivation. The desire/belief model fails to do this. The concept of normative motivation is central to our ordinary notion of acting for a reason. I argue that it is crucial in explaining the role of justifying considerations in the motivation and explanation of intentional action. I end by considering whether the approach I advocate might be extended to accommodate different versions of internalism in ethics. I suggest that the notions of normative engagement and susceptibility to the reactive attitudes, central to my account of moral judgements, can plausibly be extended to capture the kinds of intuitions that generate internalism about moral requirements and motivation.

---

## Internalism and the Problem of Moral Motivation

### 1. Introduction: The Problem of Moral Motivation

Moral judgements appear to be intimately connected with motivation and motivating states in a way that distinguishes them from standard assertions of fact. This is because judgements of fact are thought not to commit agents to acting or being motivated to act in one way rather than another. The problem of moral motivation arises because making a moral judgement does appear to commit an agent to being motivated to do or promote what is judged favourably and to refrain from or discourage what is judged unfavourably. An adequate account of moral motivation needs to explain how moral judgements, in so far as they do involve assertions of fact, can have a feature seemingly not shared by other fact-stating discourse.

A standard way of expressing the intuition that moral judgements and motivation are intimately connected is to say that moral judgements are of conceptual necessity tied to motivation. Those who support this view are said to hold an internalist account of moral judgement.<sup>1</sup> According to *moral judgement internalism* there is a connection of conceptual necessity between an agent's judgement that something is morally right and her being motivated to act on that judgement.<sup>2</sup> For example, if I see someone drop a banknote and judge that retrieving it and returning it to them is the right thing to do, this version of internalism says that I will have at least some motivation to do so. In so far as I fail to be motivated by my judgement, it will not count as a genuine moral judgement.

---

<sup>1</sup> See Falk (1948) and Frankena (1958) for early discussions of internalism. These articles primarily involve what I term below content internalism – the doctrine that there is a necessary connection between the truth of moral propositions and motivation – but are generally cited as the origins of the current debate about moral judgement internalism; see, for example, Dancy 1993: 1. This has led to a good deal of confusion which I try to address in the text.

<sup>2</sup> What we might call simple moral judgement internalism is implicit in most versions of non-cognitivism: For example, Stevenson 1937; Hare 1952: chap 5; Blackburn 1984: chap 6; Gibbard 1990: chap 1. It is criticised by Brink (1989: 45-50), and Svavarsdóttir (1999).

Internalism presents a difficulty for the cognitivist view that in making a moral judgement an agent expresses a truth-evaluable belief about how the world is morally. There is a standard non-cognitivist argument which exploits this difficulty in order to deny the fact-stating status of moral judgements.<sup>3</sup> It takes the following form:

1. Judging that an act is morally right entails motivation.
  2. No judgement that some fact or state of affairs obtains entails motivation.
- C(i) Judging that an act is morally right must be something other than a judgement that some fact or state of affairs obtains.

Premise two captures the thought that it is always *logically* possible to be indifferent to any belief or set of facts. I may judge that the truck bearing down on me will hit and kill me if I don't move out of the way, but this doesn't seem to *entail* that I am motivated to move. Nonetheless, we would expect most people to be motivated to get out of the way. This is because we think that most people have a desire for their own personal safety and thus a desire not to be hurt or killed. Notice that if I do jump out of the way of the truck, this desire not to be hurt or killed helps to explain why I get out of the way. If I desire not to be hurt or killed and believe that if the truck were to hit me it would hurt or kill me, then, it seems, I will be motivated to get out of the way of the truck.

The theory of explanation the truck example is based on is known as the desire/belief theory of action explanation and is an orthodoxy in the theory of action.<sup>4</sup> It is meant as a theory of intentional action or action done for a reason. On this model, the reason I move out of the way of the truck would be a combination of my desire not to be hurt or killed and my belief that if the truck hits me I will be hurt or killed.

Together, the desire and the belief explain why I get out of the way. This kind of

---

<sup>3</sup> The argument is not, of course, always explicit; this is a reconstruction of what I take to be its central form.

<sup>4</sup> It is also often referred to as the 'Humean theory of motivation': e.g., Dancy 1990: chap 1; Smith 1994: chap 4. Given doubts about whether Hume held such a position, I follow Schueler (1995) in calling it the desire/belief theory. See also Davidson (1980) for a version of this doctrine.

*explaining reason* is often contrasted with a good or *justifying reason* for action.<sup>5</sup> A justifying reason is a reason that counts in favour of a particular course of action. If I leap out of my seat in the cinema because I believe that the truck on the screen is going to hit me and I desire not to be hit, then this belief and desire constitute my explaining reason for jumping out of the seat. However, it seems that the reason that explains my action in this case is not a reason in favour of the action. The belief and the desire constitute my explaining reason for leaping from my seat, but what I have a justifying or good reason to do is to stay in my seat and enjoy the movie.

If the desire/belief model is correct then action that results from moral deliberation should also be explicable in terms of desire/ belief pairings. Consider Emma: Just before entering her card to withdraw money from an automatic teller machine, the machine starts to dispense fifty pound notes. Reflecting on the matter, Emma judges that the right thing to do is to return the money to the bank and notify the staff of the problem. How do we explain her action if she goes ahead and does this? According to the model above, we need to posit a relevant belief and desire pairing. Let's say that Emma believes that the right thing to do is to return the money and that she desires to do the right thing. When these states are combined, the desire/belief thesis says we have a pair of states capable of explaining Emma's returning the money.<sup>6</sup>

The crucial point about the desire/belief thesis for the issue of moral motivation is that it views desires and beliefs as distinct states neither of which can entail the other, but both of which are necessary for intentional action to take place.<sup>7</sup> The difference between the two types of states is usually put in terms of a difference in their

---

<sup>5</sup> See Brink 1989: 39-40; Schueler 1995: chap. 2. Smith (1994: chap. 4) refers to the same distinction as the 'motivating reasons/normative reasons' distinction, as does Dancy (2000: chap. 1).

<sup>6</sup> It is not always clear whether the desire/belief thesis asserts only necessary or both necessary and sufficient conditions for the explanation of intentional action (cf. Schueler 1995: 44). Smith defends the view that the model provides necessary and sufficient conditions for intentional action (1994: 92). But even if the model only provides necessary conditions for action, the problem for moral motivation raised by the non-cognitivist argument still remains.

<sup>7</sup> Cf. Smith (1994: 7): 'For any belief and desire pair that we imagine, we can always imagine someone having the desire but lacking the belief, and vice versa.' I'm not sure about this claim. While it seems *prima facie* plausible that no belief entails desire, it seems less plausible that no desire entails belief. For example, my desire that Chris stop drumming seems to entail that I believe that Chris is drumming. In the context of the present discussion, however, it is the claim that no belief entails desire that is crucial.

directions of fit (see Humberstone 1992; Smith 1994: chap. 4). Beliefs are said to be states that are out to fit the world, whereas desires are said to be states that are out to get the world to fit them. Beliefs represent the way the world is, whereas desires represent how the world is to be (Smith 1994: 7). The direction of fit of desires makes them states that are essentially or necessarily motivating. By contrast, beliefs are only capable of contributing to motivation if paired with desires (Dancy 1993: 2).

If the desire/belief thesis is correct, then no cognitive state such as belief can entail the kind of motivating or conative state such as desire that is necessary for intentional action to take place. So Emma's moral belief that she is required to return the money cannot entail that she is motivated to return the money or, indeed, that she is motivated to do what she judges to be right. It seems, then, that if we accept the desire/belief thesis we must abandon the internalist claim that moral judgements entail motivation. This is the option taken by those who defend an externalist account of the connection between moral judgements and motivation.

*Moral judgement externalism* holds that an agent's motivation to do what she judges to be morally right is independent of her judgement: it is a separate and contingent fact about her psychology that she happens to desire to do what is right.<sup>8</sup> Moral judgement internalists find this account of moral deliberation and action unacceptable because, as we have seen, they hold that morality is in some sense essentially practical or motivating. They argue that the essential practicality of morality cannot be accounted for if motivation is only contingently and externally related to moral judgements: if motivation is 'outside' rather than 'built in' to moral judgements. According to internalism, Emma's judgement that it is right to give the money back to the bank only counts as a genuine moral judgement if she is motivated in some degree to return the money. This means that once it is established that Emma has made a genuine moral judgement, there is no further question to ask about whether she is motivated to act in accordance with it. To think that there is a further question betrays, according to the internalist, some kind of *conceptual*

---

<sup>8</sup> Examples of moral judgement externalism include Brink (1989: chap 3), and Svavarsdóttir (1999). Frankena (1958) expresses externalist sympathies, as does Foot (1972).

*confusion*: it is part of what we mean when we say that someone has made a moral judgement that they are motivated to act in accordance with that judgement (cf Smith 1994: 6-7).

The problem of moral motivation, then, can be understood against the background of a particular theory of how action is motivated. With the desire/belief model in place, there is a clear tension between the view that moral judgements are fact-stating and the view that moral judgements entail motivating states such as desire. Externalists reject the latter thesis in order to defend *moral cognitivism*. This is the view that moral judgements primarily function to express beliefs about the way the world is morally and that the propositions asserted by such judgements are assessable in terms of truth and falsity. Cognitivism is a natural starting point for those wanting to defend realism about moral facts and properties, and the externalist rejection of internalism is often part of a broader defence of moral realism.<sup>9</sup> Internalists who accept the desire/belief model are likely to be non-cognitivists.<sup>10</sup> Non-cognitivism seeks to preserve the essential practicality of moral judgements by denying that such judgements express beliefs. It argues that the primary function of moral judgements is to express non-cognitive or conative attitudes. The price of adopting non-cognitivism is giving up the idea that moral judgements are primarily fact-stating. This brings with it a denial of the reality of moral facts and properties. So non-cognitivism accommodates internalism at the cost of moral cognitivism and moral realism.

Moral judgement externalism and non-cognitivism both accept the desire/belief model of action explanation. However, some theorists attempt to combine cognitivism and internalism by rejecting the desire/belief model's claim that cognitive states can only motivate with the help of desires and can never be the

---

<sup>9</sup> For example, Brink (1989: 25-29). Brink argues that we should treat moral cognitivism as the default account of moral judgements: 'Our moral judgements not only have fact-stating and property-ascribing form; they have cognitivist content as well. Many common moral judgements themselves make reference to moral properties, moral facts, moral knowledge...The form and content of our moral judgements, therefore, presupposes cognitivism' (1989: 25-26).

<sup>10</sup> See note 13, below.



source of motivation.<sup>11</sup> Cognitivist internalists argue that it is possible for beliefs to play the dominant role in motivating agents. On one version of this model, Emma's belief that returning the money is the right thing to do would be sufficient to explain her being motivated to do so. However, cognitivist internalists need not deny that desire is present in the explanation of action, only that desire rather than belief is always the source of motivation. For example, it is potentially consistent with the internalist cognitivist view of belief as the dominant motivating factor that my belief that I ought to help produces or motivates a desire to help.<sup>12</sup>

The aim of this chapter is to examine the kinds of arguments given in favour of moral judgement internalism and to distinguish it from a range of other forms of internalism in ethics with which it is often confused. In section two, I reject the non-cognitivist's use of internalism to argue against moral cognitivism and conclude that there is no quick way of dismissing the possibility of combining internalism with cognitivism. However, I accept the non-cognitivist claim that moral judgements in some sense go beyond cognition. It does seem that when an agent makes a moral judgement she is *doing* something more than merely describing some state of affairs from a disinterested point of view. Sections three and four are concerned with the task of distinguishing moral judgement internalism from other versions of internalism in ethics. In section five I try to make sense of the notion that morality is essentially practical or action-guiding and relate it to the debate between those who accept the desire/belief model of motivation and those who defend the possibility of motivating moral beliefs. I end by raising some concerns about the desire/belief model and discuss the prospects for combining internalism and cognitivism by utilising the distinction between the act of making a moral judgement and the content of a moral judgement.

The theory of moral judgement I develop in chapter three views moral judgements as linguistic or speech acts. This approach reflects the traditional non-cognitivist concern with moral language. Non-cognitivists focus on what we are doing when we

---

<sup>11</sup> Theorists who call into question the desire/belief model include Nagel 1970; McDowell 1978, 1979, 1985; Dancy 1993, 2000; Schueler 1995.

<sup>12</sup> Dancy (1993: 7-12) calls this kind of cognitivist internalism, a 'motivated-desire theory.'

call or say that something is morally right and on the point or function of moral language (e.g. Hare 1952; Stevenson 1937, 1944). This focus means that their particular account of moral psychology is based on their account of what agents are doing when they use moral language. I think the non-cognitivist emphasis on moral language is substantially correct. I argue in chapter 3 that approaching the internalism debate from the perspective of moral language rather than from that of moral psychology allows us to accommodate many of the purportedly conflicting intuitions underlying our notion of moral judgement.

Sometimes moral judgement internalism is viewed as a form of moral belief internalism: the doctrine that moral belief entails motivation (Mele 1996). However moral belief internalism is too specific a starting point for the general intuition that motivates internalism in ethics. For example, viewing moral judgement internalism as a thesis about moral belief and motivation prevents traditional non-cognitivism from being counted as a version of moral judgment internalism.<sup>13</sup> As non-cognitivism is generally held by itself and by its opponents to exemplify moral judgement internalism, it is unlikely that an account of moral judgement internalism that excludes non-cognitivism from the outset captures what we mean by this form of internalism. The point at issue in these debates is the nature of the psychological state or states involved in moral assent; moral belief internalism is simply question-begging with regard to this issue.

I think that the most fruitful and least question-begging way to think of moral judgments is in the context of their performance in speech situations. After all, it is from such contexts that we gain our experience of what it is for someone to assent to a moral proposition and our understanding of the close connection between moral judgement and action. Moral practice is a highly social and public enterprise that takes place predominately via the use of language. In chapter 5, I argue that the structure of moral psychology reflects the structure of the moral language because of the specific role language has in the development of moral institutions and practices.

---

<sup>13</sup> Non-cognitivism holds that there is no such thing as a specifically *moral* belief or propositions with specifically *moral* content. There may be a descriptive component to moral judgement and we may have beliefs about such descriptive features, but these beliefs will not be moral beliefs.



Despite this emphasis on moral language, the reader may find it useful in chapters 1 and 2 to think of moral judgements in terms mental acts or states as well in terms of linguistic acts. This in part reflects a certain ambiguity in the use of the expression ‘moral judgement’ in the literature. Most of the points I make in these chapters in favour of what I term ‘act internalism’ are not, I think, *prima facie* dependent on interpreting moral judgements as linguistic rather than as psychological acts. Nonetheless, my hope is that the survey of the debates between moral judgement internalists and externalists in these chapters will highlight the need for the linguistic turn taken in chapter three.

## 2. Non-cognitivist Internalism

Non-cognitivists argue that in order to preserve the essential practicality of morality we need to reject the view that moral judgements are fact-stating. They thus defend moral judgement internalism by giving up moral cognitivism. According to non-cognitivism, moral judgements primarily function to express conative or motivating attitudes, such as approval or disapproval, rather than to express beliefs.<sup>14</sup> On this model moral claims necessarily motivate because their primary function is to express precisely those states of an agent that are the sources of motivation.

The non-cognitivist argument for internalism relies on the purported conceptual connection between moral judgements and motivation.<sup>15</sup> Like non-cognitivism, many versions of cognitivist internalism rely on some version of this conceptual

---

<sup>14</sup> This is at least true of traditional versions of non-cognitivism, such as the emotivism of Ayer (1946) and Stevenson (1937, 1944). Hare (1952) argues that moral judgements are primarily prescriptive or command-like rather than expressive or emotive. Nonetheless, he accepts with Ayer and Stevenson that the main function of such judgements is not to state facts. Blackburn (1984) agrees that the distinctive meaning of moral terms and expressions is given by their use in expressing non-cognitive states. However, he attempts to accommodate the realist-seeming nature of ethical discourse, such as the apparently assertoric and truth-valued nature of moral judgements, by adopting a deflationary theory of truth. In this way he aims to avoid the standard objection that moral judgements on the non-cognitivist account are not assessable in terms of truth and falsity. Gibbard presents a version of non-cognitivism he terms ‘norm expressivism’. He holds that to judge an act as morally wrong is ‘to accept norms for guilt and resentment that, *prima facie*, would sanction guilt and resentment if the act were performed’ (1990: 8).

<sup>15</sup> For example: ‘It seems to be a conceptual truth that to regard something as good is to feel a pull towards promoting or choosing it, or towards wanting other people to feel the pull towards promoting or choosing it’ (Blackburn 1984: 188).

argument (e.g., Smith 1994: 6). However, the kind of explanation offered of the necessary connection between moral judgements and motivation differs considerably between non-cognitivist and cognitivist versions of internalism. Much of the appeal of non-cognitivist versions of internalism lies in their ability to explain the dynamic and practical nature of morality by appealing to the widely accepted and broadly naturalistic desire/belief model of action explanation. By contrast, cognitivist versions of internalism commonly reject the desire/belief model and replace it with controversial accounts of how cognitive states produce or cause motivation.<sup>16</sup> Some cognitivist internalists attempt to retain the desire/belief model and include motivation as part of the content of moral judgements. These theorists include subjectivists and proponents of response-dependent theories of moral value.<sup>17</sup> In both these cases, motivation, in the form of a desire or pro-attitude, is included in the truth conditions for moral propositions. Non-cognitivist internalism rejects the notion that the posited conceptual connection holds between the content of moral claims and motivation because it rejects the idea of specifically moral content.<sup>18</sup> What makes non-cognitivist internalism distinct is the view that the conceptual connection holds at the level of linguistic function rather than at the level of truth-conditional content. And because for non-cognitivists the linguistic function of moral judgements is expressive or directive, there is a straightforward conceptual connection between the *act* of making a moral judgement and motivation. Moral judgements function to allow the expression of evaluative attitudes. For the non-cognitivist, to make a moral claim without being motivated is to fail to grasp the point of moral discourse (Hare 1952: 169).

It is clear, then, that non-cognitivism interprets the internalist intuition in a distinctive way. Allan Gibbard outlines the nature of this distinctive feature of non-cognitivist internalism in the following passage:

---

<sup>16</sup> See references for footnote 11. Smith is one cognitivist internalist who does not reject the desire/belief model (1994: chap 4).

<sup>17</sup> The following theorists present 'dispositional' or response-dependent theories of value: Johnston 1989; Lewis 1989; Smith 1994.

<sup>18</sup> See, for example, Hare 1952: chap 7.

To call a thing [morally right] is to endorse it in some way. That suggests a scheme for getting at the meaning of the term. Instead of trying to define the property ["moral rightness"] by giving conditions under which a thing would have that property or lack it, start with the *use* of the term. Fix on the dictum "To call a thing [morally right] is to endorse it," and search for a sense of 'endorse' for which the dictum holds true. (Gibbard 1990: 6. My emphasis).<sup>19</sup>

The crucial feature of endorsement that relates it to internalism is its affective or non-cognitive nature: to endorse something is necessarily to hold some kind of favouring desire or pro-attitude towards the thing in question. Gibbard argues that any account of moral judgement that fails to accommodate this endorsing function, such as a descriptivist analysis, leaves a puzzle: 'It misses the chief point of calling something ['right' or 'good']: the endorsement the term connotes' (1990: 10).

For non-cognitivism, then, the endorsing or, as it is often called, commendatory function of moral judgements is the essence of the practicality of morality and is what distinguishes moral judgements from purely descriptive judgements. Another non-cognitivist, Charles Stevenson, puts the matter in the following way:

In normative ethics any description of what is the case is attended by considerations of what is to be felt and done about it; the beliefs that are in question are preparatory to guiding or redirecting attitudes. Moral judgements are concerned with recommending something for approval or disapproval; and this involves something more than a disinterested description, or a cold debate about whether it is already approved of... In this way moral judgements go beyond cognition, speaking to the conative-affective natures of men. (Stevenson 1944: 13)

The view that moral judgements have an essentially endorsing function is implicit in the non-cognitivist argument against cognitivism outlined in section one above. As we saw, the argument says that if assenting to a moral judgement entails motivation, then, because no descriptive judgement entails motivation, assenting to a moral judgement must involve something other than a descriptive judgement. But close attention to the argument shows that the premises do not support the anti-cognitivist conclusion. Rather, they support the following disjunctive conclusion:

---

<sup>19</sup> Gibbard actually talks of rationality in this passage rather than moral rightness; but I think he would hold that this point applies to all types of practical normative judgements.

C(ii) Judging that an act is morally right either involves something other than a judgement that some fact or state of affairs obtains *or something in addition to such a judgement.*

The non-cognitivist argument neglects the fact that speakers often and in the same utterance do more than just describe something: that utterances can go ‘beyond cognition’ while still being fully descriptive. If I argue that Churchill was a wartime prime minister, part of what I am *doing* is trying to convince you of the truth of the stated proposition. By contrast, if I merely state or describe the fact that Churchill was a wartime prime minister, I am in no way committed to having any desire that you accept my claim. But this contrast between arguing and describing in no way prevents a speaker from doing both. Indeed, applying the reasoning of the anti-cognitivist argument to this example generates the following *reductio*:

1. Arguing that some fact or state of affairs obtains entails that one desires the hearer to believe that the fact or state of affairs obtains.
2. No judgement that some fact or state of affairs obtains entails motivation.
- C. Arguing that some fact or state of affairs obtains must be something other than a judgment that some fact or state of affairs obtains.

So I think we should reject the non-cognitivist’s anti-cognitivist conclusion: the fact, if it is one, that moral judgements necessarily involve commendation or endorsement in no way prevents them from being genuine descriptive judgements. However, we should take seriously the non-cognitivist’s claim that moral judgements in some sense go beyond bare description. There is something about what an agent is *doing* when she makes a moral judgement that suggests she is not simply a disinterested reporter of the content of her judgement. We might say that moral judgement internalism is an expression of our interested or participatory relation to moral deliberation. Understanding precisely what this means will be a large part of the task of the rest of this work.

### 3. Content-based Arguments for Internalism

In the last section, I said that what makes non-cognitivist internalism distinct is the view that the conceptual connection between moral judgement and motivation holds at the level of linguistic function rather than at the level of truth-conditional content. The distinction between the linguistic act one performs in making a moral judgement and the content of moral judgement will be crucial to my argument in the following chapters. In this section I look at arguments for internalism which, unlike non-cognitivism, include the motivational requirement at the level of the truth conditions for moral judgements.

One way in which theorists attempt to satisfy the internalism requirement is by building motivation into the truth-conditional content of moral judgements. *Content internalism about motivation* is the view that there is a necessary connection between the truth-conditions for the propositional content of moral judgements and motivation: that is, between the truth conditions for moral propositions and motivation. On this view, part of what makes it true that an agent is morally required to perform or promote a certain act is that the agent has some motivation to perform or promote that act. It is this version of internalism that is at issue in early discussions of internalism and externalism in the literature.<sup>20</sup> Frankena borrows the labels internalist and externalist from Falk, who uses them to refer to the ‘opposition...between those who regard motivation as external and those who regard it as internal to [moral] obligation’ (Frankena 1958: 40-41).

Moral judgment internalism and content internalism about motivation are very different theses and need to be distinguished clearly from each other. This is important because moral judgement internalism and content internalism are often conflated and arguments or intuitions that support one of these positions are used to support the other.<sup>21</sup> In order to do this effectively, we need to look briefly at a

---

<sup>20</sup> Such as Falk 1948; and Frankena 1958.

<sup>21</sup> It is not always clear in Falk (1948) or in Frankena (1958) which version of internalism is under discussion, although the avowed version in both cases is content internalism (specifically, the connection between moral obligation and motivation). At least one theorist has, mistakenly I believe, interpreted Frankena’s article as concerned with moral judgement internalism rather than with content internalism: ‘Of course we should remember that Frankena’s discussion...is produced in the course of an argument about...an ‘internalistic’ conception of moral judgement, the conception according to

number of other forms of internalism in ethics. This may seem an unnecessary complication, but I think it is important to isolate moral judgement internalism from other internalist theses with which it is often confused. Moreover, while we can distinguish logically between these various internalist theses, I suspect they all stem from a much more general intuition about the connection between morality and practicality. This is why in giving an account of moral judgement internalism we should be aware of its relation to these other internalist doctrines.

Most versions of internalism are supported by the claim that implicit in our ordinary thinking about morality is the view that there is a conceptual connection between morality and motivation. For example, Mackie argues that there is a common conception of moral facts and properties as both objective and as necessarily motivating (1977: 23).<sup>22</sup> The kind of internalism Mackie is concerned with here is neither a version of moral judgement internalism nor, strictly speaking, a version of content internalism. Mackie's version of internalism involves a necessary connection between an agent's apprehension of moral facts and motivation: that is, roughly, between moral knowledge and motivation. On this conception, the motivating power of morality derives from the moral facts and properties themselves as they impact upon moral agents; and although motivation 'is a necessary consequence of perceiving or knowing [ethical facts]...motive is in no way intrinsic to ethical facts themselves'<sup>23</sup> (Darwall 1995: 10). Mackie argues that the notion of necessarily motivating properties is naturalistically 'queer' and that this gives us

---

which it is impossible for an agent to make a sincere moral judgement and not to be motivated accordingly' (Dancy 2000: 23). It is easy to see how one might gain this impression given that Frankena begins his paper by stating that it is 'concerned with a problem about the analysis of judgements of moral obligation' (1958: 40. My emphasis). But Frankena's concern is the thesis that it is impossible for someone to *have* a moral obligation without having a motivation to conform with it.

<sup>22</sup> Mackie argues that our common conception of objective values is of entities that provide 'the knower with both a direction and an *overriding* motive; something's being good both tells the person who knows this to pursue it and makes him pursue it' (1977: 40. My emphasis). Insisting that internalism asserts an overriding motivation is, I think, too strong a requirement. Unless otherwise stated, I will interpret internalism, in its various forms, as the weaker thesis that moral requirements or judgements entail *some* motivation for action (cf. Brink 1989: 41-42).

<sup>23</sup> Although the power to motivate is intrinsic to the ethical facts themselves.



good reason to reject realism in ethics (1977: 38). We might call this kind of internalism *moral knowledge internalism*.<sup>24</sup>

Moral knowledge internalism is easy to confuse with moral judgement internalism because moral judgements have a role to play with regard to moral knowledge. For example, if I know that returning the money is morally right, then presumably I judge or believe that returning the money is morally right. So, according to moral knowledge internalism true moral judgements that are knowledge will necessarily be motivating. It is clear that this doctrine is quite different to moral judgement internalism, according to which *any* genuine moral judgement, regardless of its veridicality, necessarily motivates.

Content internalism depends for its plausibility on the conceptual claim made by what I will call *constitutive existence internalism*.<sup>25</sup> Constitutive existence internalism asserts that there is a connection of conceptual necessity between moral requirements and motivation because moral requirements are in part constituted by motivation. Constitutive existence internalists, that is, endorse the view that moral requirements entail motivation because they hold that facts about moral requirements in part consist of facts about motivation. By contrast, moral knowledge internalists endorse the view that moral judgements entail motivation because they hold that facts about moral requirements necessarily bring about, but do not consist of, facts about motivation:

Though both [constitutive existence internalists and moral knowledge internalists] agree that [moral requirements] entail [motivation], they do so for opposite reasons. One takes [moral requirements] to be metaphysically basic, the other takes motivations to be... The directions of metaphysical dependence recognized are

---

<sup>24</sup> This is similar Brink's 'hybrid internalism' about motivation: '*Hybrid internalism* claims it is a conceptual truth about morality that the *recognition of a moral obligation* motivates...the agent (the person who recognises the obligation)...' (1989: 41).

<sup>25</sup> Darwall uses the label 'existence internalism' to describe this position (1995: 10). It is similar to Brink's 'agent internalism' about motivation (See Brink 1989: 40).

opposite, even though both views agree that [moral requirements] entail [motivation] and hence that not- [motivation] entails not-[moral requirement].<sup>26</sup>

The content internalist is someone who accepts constitutive existence internalism: she accepts the conceptual claim that moral requirements entail motivation because she accepts that it is part of the concept of a moral requirement that it is in part constituted by motivation. This means that part of what she asserts when she asserts that something is a moral requirement is that the motivation that partly constitutes the requirement is present: she asserts that the claim about the requirement is true partly in virtue of the claim about motivation.

For my purposes, the differences between constitutive existence internalism and content internalism are largely a matter of emphasis: the former emphasises the connection between the existence of moral facts and motivation and the latter the connection between truth conditions for moral propositions construed as asserted and motivation. I want to emphasise the latter because I am interested in the relation between moral judgements and motivation, and content internalism makes the implications of the motivational requirement for *judgements* more explicit than does constitutive existence internalism. So from here on I will talk of content internalism rather than constitutive existence internalism, while acknowledging the sense in which an acceptance of content internalism depends on an acceptance of constitutive existence internalism.

Content internalism is also easily confused with moral judgement internalism. According to content internalism, it is a condition of the truth of any moral judgement that the agent to whom it applies is motivated to act in accordance with it. On this view, then, an agent who makes a moral judgement *asserts* that a motivating state obtains. So if I judge that Emma is morally required to return the money, part of what I am asserting is that she has some motivation for returning the money. If

---

<sup>26</sup> I adapt this from Susan Hurley's excellent discussion of internalism about reasons and motivation (2001: 152). Hurley argues that most discussions of reasons internalism fail to distinguish between the logical claim that reasons entail motivation and the issue of the direction of metaphysical dependence between reasons and motivation.



this motivational condition is not part of my judgement then I have not made a genuine moral judgement, in the sense that I have failed to fully understand what it is for something to be morally right. This would be like judging that John has a new watch but it not being part of what I assert that John's new timepiece is small and worn on the body. In this case, I have not made a genuine 'watch judgement' because I have failed to understand, say, that watch judgements and clock judgements have different truth conditions. So while content internalism is a doctrine about moral propositions rather than moral judgements, it has implications for the assessment of moral judgements.<sup>27</sup> According to content internalism, my moral judgement can be genuine *even if I am not motivated by it*, as long as I judge that those to whom it applies are motivated by it. Notice that it need not be true that those to whom I direct my judgement *are* appropriately motivated. If they are not motivated, I will have still have made a genuine, although false, moral judgement. By contrast, according to moral judgement internalism, if I make a moral judgement, even about others, I must be motivated to some extent to comply with the judgement.

Moral judgement internalism is particularly difficult to distinguish from content internalism in cases where an agent takes the judgement she makes to apply to herself as well as to others. This difficulty is most pressing if we accept that an essential feature of moral judgements is their universalizability. The principle of universalizability says that if it is right for a particular agent to perform some act then it is right for any agent in relevantly similar circumstances to perform an act of the same kind.<sup>28</sup> For example, if I judge that Emma is morally required to return the money, I am committed to holding that it would be right for anyone in circumstances relevantly similar to Emma's to return the money. Given that the judgement applies to me as well as to Emma, content internalism says that I must judge that both Emma and I are, or perhaps would be under relevant conditions, motivated to comply with the requirement. So if we accept that content internalism and universalizability are correct, then my moral judgements must all make reference to my own motivations.

---

<sup>27</sup> It is these implications that Frankena is alluding to when he says his paper is concerned with an 'analysis of judgements of moral obligation' (1958: 40). He is concerned with how a particular analysis of moral obligation affects our understanding of judgements of moral obligation.

<sup>28</sup> For discussions of the principle of universalizability see Hare 1981: chap. 6; and Mackie 1977: chap. 4.

But this looks remarkably similar to moral judgement internalism: the doctrine that if I judge that it is right to  $\Phi$  then I am motivated to  $\Phi$ . The crucial, and subtle, difference is that in the case of content internalism the motivational condition is part of the *content* of the judgement: the condition to be met is that I *judge or believe* that I am motivated, not that I actually am motivated. This means that my judgement that I am morally required to  $\Phi$  may be false because my belief that I am motivated to  $\Phi$  is false. In this instance, according to content internalism my moral judgement would still be genuine. By contrast, as we have seen, moral judgement internalism requires that I have an *actual motivation* to do what I judge to be right in order for my judgement to be genuine.

At this point it is worth noting that content internalism comes in counterfactual as well as occurrent versions. It may be the case, for example, that neither Emma nor I is well placed to consider the moral implications of her not returning the money, and so neither of us is presently motivated to return the money. This need not mean that the ascription of moral obligation to us must be withdrawn. The counterfactual version of content internalism says that the moral obligation will still apply to Emma and to me if, were we to consider the situation from an appropriate point of view or under relevant conditions, we would both be motivated to return the money (cf. Copp 1995: 189-190). This version of content internalism presents no problem for my position: it merely emphasises the differences I have been stressing between moral judgement internalism and content internalism. If we accept the view that moral judgement internalism is a doctrine about the connection between moral judgements and occurrent motivation, then counterfactual content internalism, by its very nature, fails to satisfy this requirement. So one important difference between both types of content internalism and moral judgement internalism is that the former fail to guarantee the necessary connection between moral judgements and occurrent motivation asserted by the latter. More than this, however, both forms of content internalism fail to capture the sense in which, on the non-cognitivist model of internalism, moral judgements function to allow the *expression* and not the assertion of motivating attitudes.

Given these differences between moral judgement internalism and content internalism, we might wonder why theorists have thought that the intuitions which support one of these positions can be used to support the other.<sup>29</sup> One reason is the generality of the internalist intuition. If we read the internalist's claim about moral judgements to be merely that there is a conceptual connection between moral judgement and motivation, then content internalism can plausibly be seen as satisfying this claim. This is because according to content internalism there is a straightforward conceptual connection between the content of moral judgements and motivation. Part of the problem here is that the expression 'moral judgement' is structurally ambiguous: it can refer either to the content of the judgement, the moral proposition, or to the act of making the judgement. This is the difference between act and content alluded to in the previous section. Although a simple distinction, I believe attending to it will allow us to become much clearer about what is at issue between internalists and externalists in ethics, and to better understand the problem of moral motivation. Remember that non-cognitivism views internalism as a thesis about judgement acts rather than judgement content. This is the positive part of the non-cognitivist thesis. I have already argued that we should reject the negative or critical part of non-cognitivism. Given that it is this negative part of non-cognitivism that gives it its title, we need another title for the kind of internalism espoused by non-cognitivism. I will call the view that moral judgement internalism is a thesis about judgement acts *act internalism*. In the rest of this section I give reasons why we should prefer act internalism to content internalism as an interpretation of moral judgement internalism.

The most obvious reason for interpreting moral judgement internalism as a doctrine about judgement acts rather than contents is that content internalism is only incidentally about moral judgements. As we have seen, content internalism is really about moral propositions or moral facts: it has implications for moral judgements but

---

<sup>29</sup> Smith's defence of his 'practicality requirement' appears to assume that some form of content internalism can establish a version of moral judgement internalism (1994: chap. 3). I make the point that the practicality requirement relies on a version of content internalism in my discussion of Smith in chap. 2.4 above. Copp's discussion of moral judgement internalism (what he calls 'belief internalism') also implies that content internalism might support a version of moral judgement internalism (1995: 190-192).

is not a doctrine about moral judgements. By contrast, act internalism is essentially about the connection between moral judgements and motivation. The nature of content internalism also means that it does not secure the necessary connection between moral judgement and motivation thought to be indicative of moral judgement internalism. We saw above that content internalism only secures a necessary connection between moral judgements and *judgements* about motivation and not between moral judgements and actual motivation. Another weakness of content internalism is that it fails to accommodate the endorsing or commendatory nature of moral judgement. According to content internalism, what one is doing when making a moral judgement is *describing* or *asserting* the fact that one is motivated in the appropriate way. If one's judgement is true, then one is indeed motivated to perform the act; but one's judgement may be false, in which case one may have no desire to do what one judges it is right to do. Moreover, it is not clear that in simply describing the fact that one is motivated, even if it is true, one endorses or commends whatever it is that one is motivated to do.

Perhaps the most compelling reason for not interpreting moral judgement internalism as a form of content internalism is that if we do so it becomes more difficult to defend the status of internalism as a platitude or commonplace about the nature of morality. This is because act internalism is more intuitively compelling than content internalism. It will help here to look at the standard kind of example designed to show the intuitive appeal of internalism. Suppose I tell you that we all have a moral obligation to vote in general elections. Given this information about me you form expectations about my behaviour. Specifically, you expect me to go and vote when it comes to the next general election, or, if I don't, for there to be a good reason for my not voting. So when the day of the next general election arrives and I don't vote, you ask me to explain myself. I respond that while I genuinely think that I have a moral obligation to vote, I am in no way, and have never been, motivated to vote. The standard internalist response is to say that because I have failed to be motivated there is some sense in which I have not made a genuine moral judgement.<sup>30</sup> The question is whether my mistake is to fail to recognise that in order for something to

---

<sup>30</sup> See my chap. 2 for a discussion of Hare's 'inverted commas' version of this argument.

count as a moral requirement for me I must be motivated to adhere to it or whether my mistake is of some other kind.

The content internalist must explain my failure to make a genuine moral judgement in terms of my not understanding that motivation is one of the truth conditions for having a moral requirement. It is this rather than the fact that I am not motivated which constitutes the failure of my judgement. If I had merely believed mistakenly that I was motivated to vote my judgement would still be sound. But this seems an unlikely interpretation of the thought experiment because the view that moral requirements are contingent on an agent's desires is simply not intuitively compelling in the way internalism is meant to be. In fact it runs counter to the influential Kantian doctrine that moral requirements are categorical imperatives which apply to agents regardless of their contingent wants (cf. Brink 1989: 45; Dancy 1993: 4.). We can test this Kantian intuition by extending the example. Suppose you agree with my claim that we all have a moral obligation to vote in general elections. It seems unlikely that once you discover that I am not in fact motivated to vote you would withdraw the requirement from me.<sup>31</sup> This suggests that the internalist intuition, if it is a platitude, must be something other than the content internalist's claim that moral requirements are contingent on agents' desires.

My proposal is that interpreting the internalist intuition along the lines of act internalism provides a more satisfactory interpretation of thought experiments like the one above. According to act internalism, not being motivated by one's moral judgements is like claiming to argue that something is the case but having no desire to convince the hearer of the truth of one's claim. If I argue sincerely that you are wrong, part of what I am doing is attempting to convince you of your mistake. Similarly, if I judge that something is morally required, part of what I am doing is commending or endorsing that option. It seems clear that if I try to convince you that

---

<sup>31</sup> The counterfactual version of content internalism might seem to fare better than the occurrent version in dealing with the issue of categoricity. However, consider the moral obligation not to torture people for fun. Let's say Jane is in whatever ideal conditions the counterfactual theorist posits, but that she still fails to have any motivation to refrain from torturing for fun. I suggest it is implausible that we would withdraw from Jane the requirement not to torture for fun simply because she has no desire not to torture people.

p, then I desire that you accept that p. So we can say that there is a necessary connection between a speaker arguing that p and desiring the hearer to accept that p. Act internalism implies that if I commend an option then I desire that option to be taken. We could then say that there is a necessary connection between judging that an option is morally required and desiring that option to be taken and thus being motivated to take or encourage the taking of that option. So act internalism satisfies moral judgement internalism: the thesis that if an agent judges that something is morally right, then she is motivated to act in accordance with her judgement. Moreover, as act internalism does not assert a necessary connection between moral requirements and motivation or desire, it satisfies moral judgement internalism while being consistent with the Kantian doctrine that moral requirements are categorical imperatives.

Although act internalism provides a better interpretation of moral judgement internalism than does content internalism, a number of difficulties still remain. The way the voting example was presented above didn't distinguish between commending some act and being motivated to perform or promote such an act. We simply assumed that if an agent commends something then she is motivated to do it. Indeed, this is the way commendation is generally treated by non-cognitivists. But as will become clear in following chapters, there are cases in which failures of motivation are explicable and perhaps to be expected.<sup>32</sup> In such cases an agent still seems to be commending the act he judges as morally right, even though he fails to be motivated to perform it. Another issue, dealt with in the next section, is that the incoherence attributed to the person who fails to be influenced by her moral judgement is sometimes said to result from her failing to take the judgement as providing a reason or justification to perform the act in question. These issues complicate matters but do not, I think, undermine the arguments for preferring act internalism to content internalism as an interpretation of the intuitions underlying moral judgement internalism.

---

<sup>32</sup> See especially chap. 4.2.



#### 4. Internalism and Justifying Reasons

In the last section, I stressed the importance of distinguishing between content internalism, which primarily concerns moral propositions, and moral judgement internalism proper, and I noted that the distinction is often not clearly observed. Another crucial distinction in the internalism debate is that between internalism about motivating states and internalism about practical justification or justifying reasons. We have already discussed content internalism and judgement internalism about motivating states. Below I discuss how these positions relate to content internalism and judgement internalism about practical justification.

The four main elements to keep in mind when discussing internalism in ethics are moral judgements, moral content, motivating states, and practical justification (Brink 1989: chap. 3). The different versions of internalism in ethics can be understood in terms of how the former two elements are combined with the latter two elements. Moral judgement internalism can be about either motivating states, as above, or about justifying reasons. Moral judgement internalism about motivating states says that there is a necessary connection between moral judgements and motivating states; moral judgement internalism about justifying reasons says that there is a necessary connection between moral judgements and justifying reasons. Content internalism can also be about either motivating states or about justifying reasons. Content internalism about motivating states says that there is a necessary connection between the truth of moral propositions and motivating states; content internalism about justifying reasons says that there is a necessary connection between the truth of moral propositions and justifying reasons.

To get some idea of how discussions of internalism often run these elements together it will be useful to look at Jonathan Dancy's account of Nagel on internalism. Dancy begins with the following passage from Nagel:

Internalism is the view that the presence of a motivation for acting morally is guaranteed by the truth of the ethical propositions themselves. On this view the motivation must be so tied to the meaning, or truth, of ethical statements that when in a particular case someone is (or perhaps merely believes that he is) morally required to do something, it follows that he has a motivation for doing it. Externalism holds, on the other hand, that the necessary motivation

is not supplied by ethical principles and judgements themselves, and that an additional psychological sanction is required to motivate our compliance...Internalism's appeal derives from the conviction that one cannot accept or assert sincerely any ethical proposition without accepting at least a prima facie motivation for action in accordance with it. (Nagel 1970: 7)

Dancy then makes the following observation about the passage:

Nagel's work here is pioneering but, like many other pioneers, he did not find the best road for others to follow. There are two ideas which he fails to keep apart in this passage. The first is that motivation is provided by the mere truth of some proposition; the second is that the motivation is provided by one's belief in that proposition, whether it be true or not. The second of these is much less outlandish than the first, and is what I take to have been Nagel's intention. (Dancy 1993: 1)

In fact there are more than two ideas that Nagel fails to keep apart, and because Dancy doesn't recognise this he falls into similar difficulties to Nagel. Dancy is correct in pointing out that Nagel shifts from talking about moral truth to talking about moral belief or moral judgements. Nagel implies that internalism about moral truth and internalism about moral judgements are equivalent. This is why he says that if 'someone is (*or perhaps merely believes that he is*) morally required to do something, it follows that he has a motivation for doing it.' But if, as he says, motivation is guaranteed by the truth or content of an ethical proposition, then it does not follow that a moral judgement or belief with this content will entail motivation. This is because it is always possible that the judgement or belief is false and that the motivation condition is not met.

Dancy asserts that the notion that motivation is connected to the truth of ethical propositions is 'outlandish'. It is worth reflecting on why Dancy might think this given that Nagel makes it explicit throughout the passage that this is how he construes the thesis of internalism. One reason Dancy might think this view outlandish is because it suggests a form of content internalism about motivation and such a view contradicts the notion of moral requirements as categorical imperatives by making moral truth depend on contingent desires (cf. Dancy 1993: 4). But Nagel is clearly not a content internalist in this sense. He precisely does not want moral requirements to depend on desires in this way:



The position which I shall defend... provides an account of ethical motivation which does not rely on the assumption that a motivational factor is already present among the conditions of any moral requirement. On this view the possibility of appropriate motivation must be guaranteed by the truth of the moral claim itself – *but not because the existence of such motivation is included in advance among the independently comprehensible truth conditions of every moral claim.* (Nagel 1970: 13. My emphasis)<sup>33</sup>

How else might Dancy interpret the notion of motivation being provided by the truth of a moral proposition? He might view it as the claim that the truth of an ethical proposition has the power to influence agents regardless of whether or not they are aware of it. This does seem an extreme view, but it can be made more plausible if changed to the claim that moral truths necessarily motivate agents but only when they become aware of them. This is the view I called moral knowledge internalism. But this doesn't quite capture what Nagel is saying. Nagel does seem to want to say that all sincere moral judgements entail that one accepts a motivation for action. And this would not be the case if only those moral judgements that were true entailed motivation, as there would potentially be many sincerely held but false moral judgements that did not motivate.

In order to make sense of Nagel we need to recognise that, in addition to the confusion about moral truth and moral judgement noted by Dancy, he doesn't clearly distinguish between internalism as a thesis about motivation and internalism as a thesis about justifying reasons. If we attend to all these distinctions it becomes clear that much of what Nagel says suggests a version of internalism that asserts a necessary connection between moral requirements and *justifying* reasons (or motives) for action. Contrary to Dancy's claim, it is clear that Nagel interprets his thesis as concerned primarily with the truth of moral propositions and only incidentally with moral judgements. And this emphasis would not fit with Nagel's broader position if the posited necessary connection were said to hold between moral truth and *motivating states*. By contrast, the view that moral requirements entail justifying reasons for action is perfectly consistent with Nagel's theory. The last sentence of the passage, at least, clearly suggests that Nagel has in mind the notion of motivation as a justifying motive for action. If it is part of the truth-conditional content of moral

---

<sup>33</sup> Nagel continues: 'There are reasons for action that are specifically moral; it is because they represent moral requirements that they can motivate, and not vice versa' (1970: 13).

judgements that moral requirements provide justifying motives for action, then we can make sense of the notion that to make a sincere moral judgement one must accept 'a prima facie motivation for action.' By contrast, it is unclear what it might mean to *accept* a prima facie motivating state for action. So I think that at least part of what Nagel is arguing for under the banner of internalism is the view that moral requirements necessarily provide justifying motives or reasons for action (cf. Brink 1989: 51).

We need to be careful here not to downplay the extent to which Nagel sees himself as providing an account of the motivational efficacy of moral judgements as well as their justifying force. Strictly speaking, *the only kind of internalism that Nagel supports is content internalism about justifying reasons*: the thesis that moral requirements entail justifying reasons for action. However, he takes it to be obvious that a theory of practical justification must mesh with an account of motivational influence, and he rejects the separation of motivational from normative discourse (Nagel 1970:15, 28). So in addition to arguing for content internalism about justifying reasons, Nagel wants to defend the notion that justifying reasons are capable in their own right of having a motivational effect on agents. As Nagel sees it, the desire/belief theory of motivation blocks this possibility because it denies that mere knowledge or belief can motivate in the absence of desire or be the dominant force in motivation. So he takes it that his defence of content internalism about justifying reasons needs to be supplemented by a cognitivist account of moral motivation. This allows him to provide an account of the motivating force of moral considerations in terms of their justifying force. Nagel refers to this connection between justification and motivation in the following passage:

Kant's effort to produce a categorical imperative is an attempt to discover requirements on action which apply to a man on no conditions about what he wants, how he feels, etc. *They must nevertheless be requirements whose validity involves the capacity to be motivated in accordance with them.* Since this motivational factor cannot come from a presupposed motivation which is made a condition of the requirements, it must...come from the requirements themselves. That is, what makes the requirements valid for us must itself determine the capacity of our motivational structure to yield corresponding action. (Nagel 1970: 12. My emphasis.)

However, even though Nagel supports a connection between content internalism and the theory of motivation, he does not support a necessary connection between moral judgements and motivation and thus does not support moral judgement internalism:

Hence I conclude that the judgement that one has a reason to do something includes the acceptance of a justification for doing it, and that this is its motivational content...[however] motivational content does not necessarily imply motivational efficacy. (Nagel 1970: 65)

Dancy's mistake is to read Nagel's cognitivist theory of motivation as a form of moral judgement or moral belief internalism and to identify Nagel's internalism with this doctrine and not the form of internalism Nagel does hold, namely, content internalism about justifying reasons. At times Dancy, himself, is unclear about the doctrine he is defending. At one point he describes cognitivist internalism as the doctrine that moral judgements express beliefs 'which cannot be present without motivating.' (Dancy 1993: 3). A few lines further on, however, internalism is associated with the doctrine that one cannot make a moral judgement without seeing it as practically relevant, which suggests content internalism about practical justification.<sup>34</sup> But Dancy has already rejected interpreting internalism as a doctrine about moral content. So it is surprising when his main argument for internalism is actually an argument for moral content internalism about justifying reasons and not for moral judgement internalism:

[T]here must strong intuitive reasons in favour of internalism... And there is at least one such reason, namely the sense that morality is essentially practical, so that it would be odd for someone to say 'This action is wrong but I don't see that as at all relevant to my choice'... [This] can be backed up by the thought that moral considerations are ones whose practical relevance cannot be escaped by saying 'I don't care about that sort of thing.' (Dancy 1993: 4)

The position this argument defends is that moral requirements entail justifying reasons for action which apply to agents independently of their desires. This doctrine entails that if someone judges that something is morally required then part of what they judge is that doing that thing is practically justified. It does not entail

---

<sup>34</sup> As it turns out, Dancy appears to opt for a position somewhat similar to the one I attribute to Nagel. In order to accommodate the problem of accidie or listlessness, Dancy rejects a strict entailment between moral beliefs and motivation and adopts a defeasible motivation theory according to which moral beliefs are capable of motivating in their own right but can be present without motivating (see Dancy 1993: 24).

that if someone judges something is morally required they express a belief which cannot be present without motivating. It is perfectly compatible with content internalism about justifying reasons that an agent might recognise a moral requirement and its practical significance and remain unmoved by it. As I said, it is to ameliorate, if not remove, this possibility that Nagel develops his cognitivist theory of motivation. But there is an immense difference between establishing that certain beliefs are *capable* of motivating independently of desires and establishing that certain beliefs necessarily motivate independently of desires. So Dancy's argument for internalism is not the argument for a cognitivist version of moral judgement internalism that he aimed to provide.

In this section, I have argued that just as theorists discussing internalism often fail to distinguish between moral content and moral judgements, so they often fail to distinguish between motivating states and justifying reasons. It is because Nagel does not distinguish clearly between motivating states and justifying reasons that Dancy interprets him as propounding some form of moral judgement internalism. However, Nagel in fact advances a form of content internalism about justifying reasons and not a form of moral judgement internalism. It sometimes seems as if Nagel supports moral judgement internalism because he supplements his content internalism with a cognitivist account of moral motivation. But the cognitivist account of moral motivation does not assert a necessary connection between moral judgements and motivation. Rather, Nagel asserts the weaker claim that agents *can* be, and normally are, motivated by their moral judgements because such judgements express beliefs with motivational content capable of moving them to act independently of prior, independent desires.<sup>35</sup>

Even though I said that Nagel's *internalism* is primarily about moral content rather than moral judgements or beliefs, his cognitivist account of moral motivation plays a central role in vindicating his internalism. Nagel is aware that the issues underlying

---

<sup>35</sup> There is some debate about how best to interpret the role of desire in Nagel's theory. In *Moral Reasons* (1993), Dancy argues that, on Nagel's view, moral beliefs produce robust desires that help to explain intentional action. However, in *Practical Reality* (2000) he adopts the view that Nagel's desires are only consequentially ascribed to agents insofar as their action is deemed intentional. See chap. 2.4 (iii) for a more detailed discussion of these issues.

the problem of moral motivation are simply deferred if all one does is establish that moral judgements are judgements about what is practically justified. This is because there can be internalism and externalism about judgements of practical justification and motivation as well as about moral judgements. Internalists about practical judgements assert that if an agent judges that something is practically justified then she is motivated to act on her judgement.<sup>36</sup> Externalists say that an agent must have a prior, independent desire, such as a desire to do what is practically justified, in order to be motivated to act on her judgement. Now, Nagel is strictly speaking not an internalist about practical judgements and motivation because he doesn't hold that practical judgements *entail* motivation. However, neither is he an externalist, as he thinks that practical judgements and beliefs can motivate without the support of prior, independent desires. Nagel holds a cognitivist theory of motivation according to which practical judgements can, and usually do, motivate agents to do what they judge to be practically required. And because moral judgements are a form of practical judgement, they too have the capacity to motivate independently of prior, independent desires.

Like content internalism about motivation, Nagel's approach fails to guarantee a necessary connection between moral judgement and motivation and so fails to satisfy moral judgement internalism. Importantly, however, unlike content internalism about motivation Nagel defends the notion of motivating moral beliefs and so provides an alternative to moral judgement externalism. His approach also respects the Kantian notion of moral requirements as categorical imperatives. However, his theory doesn't appear to account for the endorsing or commendatory function that seems integral to the act of making a moral judgement. And we might worry that a failure to secure a necessary connection between moral judgements and motivation leaves it unexplained why and when agents are motivated by their moral beliefs. We need a more detailed account of how content internalism about justifying reasons combines with the cognitive theory of motivation to explain the fact that agents are usually motivated to do what they judge to be morally right.

---

<sup>36</sup> Williams (1980) argues that motivation is internal to justifying reasons.



## 5. Action-guiding Content and Motivating States

Internalism in ethics is often defended on the grounds that it is necessary in order to secure the essential practicality of morality. We have seen in this chapter that there is not one but a range of internalist positions in ethics. These different positions attempt to capture the various ways in which morality is practical. Two forms of internalism in particular seem necessary in this respect. The form of internalism espoused by non-cognitivism, which I have called act internalism, captures the *motivating force* of moral judgements: moral judgements express just those states that are the sources of motivation. Content internalism about justifying reasons, by contrast, captures the *justifying authority* of the propositional content of moral judgements. Thus we can divide the claim that morality is essentially practical into two parts. The first part is that moral requirements are essentially practically justifying and the second part that moral judgements are essentially motivating. If morality is essentially practical in these two ways then both ways must be captured by an adequate theory of moral judgement.

Non-cognitivists often claim that their theory accommodates the practicality of morality in a way moral cognitivism cannot. Hare, for example, argues that the function of moral principles is to guide conduct and that his version of non-cognitivism captures this action-guiding or practical nature of morality (1952: 1). However, while non-cognitivism plausibly accommodates the motivating force of moral judgements, it is not clear that it can accommodate the justifying authority of moral requirements. Strictly speaking, non-cognitivism denies that there are facts about moral requirements and so denies that there is such a thing as moral authority. Of course, sophisticated versions of non-cognitivism attempt to account for the realist seeming nature of moral discourse.<sup>37</sup> So non-cognitivism can account for the way in which moral requirements seem justifying; but this is hardly the same as giving an account of how moral requirements might actually be practically justifying. Consequently, non-cognitivists cannot easily claim that their theory accounts for the action-guiding character of morality.

---

<sup>37</sup> I am thinking here of Blackburn's quasi-realism (1984: chap 6).

When we talk of the action-guiding character of morality we are talking both of its normative or justifying force and the role that this practical normativity plays in our deliberation about how to act. If the point of moral discourse is to guide conduct, then moral judgements must have action-guiding content: content which tells the agent that she is practically required or has a justified reason to behave in a certain way. And if moral norms play a genuine role in deliberation about what to do, then it must be possible that agents are sometimes motivated by the normative content of their moral judgements. This suggests that only some form of cognitivism can capture the action-guiding character of moral judgements.

In section two, we saw that the standard non-cognitivist argument against cognitivism fails. This failure suggests that there is no quick way of dismissing the union of cognitivism and internalism. But some attempts to combine cognitivism and internalism fare better than others. The version of cognitivist internalism that includes motivation in the truth conditions for moral requirements marries content internalism about motivation with cognitivism. However, content internalism does not have the same intuitive appeal as moral judgement internalism and cannot secure the connection between moral judgements and motivation that moral judgement internalism asserts. Nagel's version of internalism avoids the counterintuitive nature of content internalism about motivation by construing internalism as a thesis about the relation between moral requirements and justifying reasons. But Nagel's position does not support moral judgement internalism: he acknowledges that there is no strict entailment relation between moral judgements and motivation. So while there is no quick argument against cognitivist internalism, it does seem difficult to develop a form of moral cognitivism which accommodates a necessary connection between moral judgements and motivation.

We need a better understanding of why the notion of necessarily motivating cognitive states proves so difficult to accommodate. One reason might be that cognitive states are simply not in the business of motivating. It is often said that cognitive states have a different 'direction of fit' to characteristically motivating states like desire. Dancy explains:



Desires are states which are guaranteed to motivate; they cannot exist without motivating. We can say that they are essentially or necessarily motivating states. Belief, on the other hand, requires the help of desire if it is to motivate... One way of putting this point would appeal to the notion of direction of fit. Beliefs are supposed to fit the world; they have the mind-to-world direction of fit. Desires try to get the world to fit them when it doesn't need to; they have the world-to-mind direction of fit. Given this picture...desires must be internally motivating states, since their essence is to try to get the world to fit them. (Dancy 1993: 3)

The notion of desires and beliefs having different directions of fit supports the desire/belief model of action explanation outlined in section one. This is because if desires and beliefs perform different but complementary tasks, both states will be needed in order for intentional action to take place. Thus it is often said that for an agent to be motivated to act there must be some end that an agent wants and a belief or beliefs about how to achieve that end. I leave a more detailed discussion of the desire/belief theory until chapter five; however, there are a couple of points it will be useful to keep in mind. The first point is that discussions of the desire/belief model rarely distinguish between it as an account of action explanation and as an account of motivation. These are quite distinct issues and need to be kept apart. For example, it may be that the desire/belief model provides the correct account of action-explanation but not the correct account of what motivates the ends towards which action is directed. A second and related point is that the term 'motivation' is highly ambiguous. This causes much confusion in the literature on moral motivation. As I see it, 'motivation' can refer either to the psychological state of being motivated, to the consideration in light of which an agent is motivated, or to the psychological state which conveys the consideration in light of which the agent is motivated. The first of these would typically be called a desire, the second the agent's reason, and the third a justifying belief, with a reason claim as its content. While we might accept readily that beliefs cannot be motivating states in the first sense given their direction of fit, this doesn't seem to prevent them from being motivating in the sense of conveying the consideration in light of which an agent is motivated to act. In fact, it seems this task would require a state with the direction-of-fit of a belief. I leave these difficult issues for the moment with the understanding that they will have to be addressed at a later point.

## 6. Accommodating the Internalist Intuition

In this chapter I have distinguished moral judgement internalism from a range of other internalist doctrines in ethics. Specifically, I have argued that moral judgement internalism should be viewed as a doctrine about the concept of the act of making a moral judgement. However, at a number of points it has been suggested that even this formulation is too vague to elucidate sufficiently clearly the range of intuitions underlying moral judgement internalism. In this section I want to clarify which (moral judgement) internalist intuitions I think a theory of moral judgement needs to accommodate and which interpretations of internalism it would be best to abandon.

In what follows I assume the act internalist interpretation of moral judgement internalism. In very general terms this version of internalism says that moral judgements are of conceptual necessity tied to motivation. This view is commonly interpreted as asserting a connection of conceptual necessity between an agent's judgement that something is morally right and her being motivated to act on that judgement. We might call this version of internalism, simple moral judgement internalism. Consider Smith's formulation of simple moral judgement internalism:

If an agent judges that it is right for her to  $\Phi$  in circumstances C, then she is motivated to  $\Phi$  in C. (Smith 1994: 61)

Smith rejects simple moral judgement internalism because it fails to account for cases of weakness of will and mood-based motivational failures. I discuss these cases in chapter 4.2. I agree with Smith that there do seem to be instances of agents making genuine moral judgements and failing to be motivated as a result of psychological disorders of various sorts. This means I also reject simple moral judgement internalism.<sup>38</sup>

It seems clear that as a result of severe depression or despair an agent may have no desire to do the particular act prescribed by her moral judgement. Suppose John

---

<sup>38</sup> Given Smith's adherence to the desire/belief thesis, I assume that his statement of internalism is equivalent to the following formulation: If an agent judges that it is right for her to  $\Phi$  in circumstances C, then she desires to  $\Phi$  in C.

judges that he morally ought to vote in the national election but due to high levels of anxiety and depression cannot bring himself to leave the house. It seems odd to say that he can no longer hold genuine or sincere views about what he morally ought to do, views he held prior to his depression, simply because of a failure of motivation. The difficulty with such examples is that our intuitions pull us in internalist and externalist directions: we think that agents can fail to want to do what they genuinely judge they morally ought to do, but we also think that an agent is not genuine or sincere in her moral judgement if she is not committed to it in some way that goes beyond bare cognition or belief. I propose to address this tension by way of the difference, mentioned at the end of section 1.3, between endorsing the performance of some act and being motivated to perform or promote the act, where ‘being motivated’ is read as desiring to perform or promote the act. I suggest that the most compelling and plausible version of moral judgement internalism asserts a connection of conceptual necessity between judging that an act is morally right and endorsing or commending the performance of that act. This gives us the following formulation:

If an agent judges that it is morally right for her to  $\Phi$  in circumstances  $C$ , then she commends (endorses, approves)  $\Phi$ -ing in  $C$ .

Recall that this is the view of internalism espoused by non-cognitivism. Non-cognitivists argue that moral judgements have an essentially endorsing or commending function. It is this reading of internalism that I aim to accommodate in the rest of this work. Unlike simple versions of non-cognitivism, I do not think that to commend some act is necessarily to be motivated to perform the act, in the sense of having a desire to perform the act. There are cases in which an agent commends some act as morally right but fails to have any desire to perform the act.

Nonetheless, commending something as right necessarily involves doing something that goes beyond bare cognition in a way that is essentially connected to the conative-affective natures of agents (cf. Stevenson 1944: 13).

I said above that I agree with Smith's rejection of the doctrine that a genuine judgement of moral rightness entails a desire to do the act prescribed by the judgement. There are a number of reasons to reject this view of internalism. One reason is that, as stated, it does not easily accommodate the distinction between sincere and insincere moral judgements, viewed as linguistic acts. As I argue in chapter 3.5, if an agent makes an insincere moral judgement, she still expresses or purports to have the psychological state associated with sincere examples of such judgements. Provided the agent's judgement meets other relevant 'success conditions' we should count such insincere moral utterances as genuine instances of moral judgements.<sup>39</sup> This difficulty is easily met by making internalism a thesis about sincere moral judgements and motivation:

If an agent judges **sincerely** that it is right for her to  $\Phi$  in circumstances C, then she desires to  $\Phi$  in C.

However, this version of internalism still fails to accommodate the kinds of failures of motivation that can result from depression and other sorts of mood disorders. This difficulty leads Smith to reject this version of internalism as too strong (1994: 61). While there is a sense in which the entailment claim it makes is too strong, there is also a sense in which it expresses quite a weak motivational requirement: it does not require the agent to in any way endorse the act or to be motivated by the judged rightness of the act. For example, suppose that John believes it is right to give money to charity but fails to have any desire to do so until someone convinces him that giving money to charity will make others think better of him. Now John believes that it is right to give money to charity and desires to give money to charity, so satisfying the internalist requirement. But the presence of this desire does not seem appropriately relevant to the genuineness of John's moral stance. It does not seem to capture the robust sense of endorsement or commendation that I have argued

---

<sup>39</sup> This is a point about moral judgements viewed as linguistic or speech acts. In general it is only in the context of performing linguistic acts that an agent can issue insincere judgements, promises, assertions, commands, etc. In 3.5, I argue that we should allow insincere speech acts to count as genuine, albeit defective, examples of linguistic acts. Denying this status to insincere speech acts would mean, for example, that in the case of promises we could not hold someone accountable if they

is central to the function of moral judgements. While it is true that internalism states a necessary and not a sufficient condition for genuine moral judgements, if it is not to be empty as a claim about morality, it must state a motivational condition that is clearly related to the status of such judgements.

A fundamental weakness, then, with the simple moral judgement internalism is that it does not account for the intuitive distinction between merely desiring to perform an act and endorsing the performance of an act. For example, it would allow an agent to believe that the morally right thing to do is to stop torturing the dog and to desire to stop torturing the dog, but not to endorse stopping the torture of the dog. Moral judgements go beyond belief in involving endorsing attitudes, but they also go beyond desire. I can have a strong desire and thus be motivated to eat the rest of the birthday cake without in any way endorsing or commending my eating the rest of the cake. To desire to perform an act is not, it would seem, the same as commending or endorsing the performance of the act. An adequate account of moral judgements needs to make sense of this distinction. I address these issues in chapters 4 and 5.

Any plausible account of moral judgement internalism must help to elucidate the nature of moral motivation. Consequently, my favoured interpretation of internalism needs to provide an account of the connection between commendation and motivation. Early versions of non-cognitivism suggest that to commend, endorse or approve of some option is simply to desire that option. If we accept the claim that desire is an internally motivating state, then such versions of non-cognitivism help to explain how moral judgements motivate. As we have seen, however, such a view struggles to account for the problem of moral motivational failure and the problem of unendorsed desires. The account of moral judgement I propose aims to capture the sense in which genuine and sincere moral judgements entail an agent's endorsement of the option that is judged to be right and the sense in which this notion of endorsement goes beyond bare belief or cognition in linking up with the kinds of conative states standardly associated with action explanation and motivation.

---

were insincere, as they would not have made a promise. Similarly, we could not easily speak of

There is a sense in which the position I develop does support the view that moral judgements entail motivation. I argue in chapter 4 that in order to make a genuine and sincere moral judgement an agent must be motivationally or conatively disposed to be guided by the content of her moral deliberations. Genuine moral judgements assume a background of affective-conative engagement with normative practice. There is thus a connection of conceptual necessity between moral judgements and the range of conative attitudes that make up the stance of the normatively engaged agent.<sup>40</sup> These attitudes are sometimes termed moral emotions or reactive attitudes and include resentment, guilt, shame, moral indignation, and moral approval. This position is clearly distinct from the externalist view that the possession of a relevant moral belief is sufficient for an agent to make a genuine and sincere moral judgement. The kind of attitude I associate with moral commitment also secures an internal connection between moral judgements and conative states. In chapter 4, I argue that an agent is committed to the requirement expressed by her moral judgement if she 'normatively expects' herself and all relevant others to be guided by that requirement. To hold oneself and others to a requirement in this way is to be disposed to respond to violations of the requirement with negative sanctioning or reactive attitudes, such as guilt, resentment or moral indignation. This aspect of my account of moral judgements also distinguishes it clearly from standard version of moral judgement externalism.

It is worth keeping in mind that it is not what a position is called but what it amounts to that is crucial in these debates. If the position I advocate excludes certain implausible versions of internalism and accommodates some powerful intuitions behind externalism, or at least assuages some externalist worries about internalism, then we should see this as a virtue rather than a weakness of my position. The debate between moral judgement internalists and externalists is typical of a tendency in philosophy for certain very narrow versions of doctrines to acquire a life of their own in the literature and to become divorced from the general intuitions which first

---

insincere statements or insincere requests, as such speech acts would either be sincere or not made.  
<sup>40</sup> The conceptual connection is at one remove, being mediated by the concept of endorsement or commendation. In chapter 4, I argue that we should think of the kind of endorsement involved in moral judgements in terms of the notion of 'holding to account' or 'normative expectation.'



inspired them. One of my aims is to capture what I take to be the general and persistent intuition that genuine moral judgements are necessarily tied to some kind of conative or practical inclination towards those acts judged favourably and against those acts judged unfavourably.

The version of moral judgement internalism I defend is a strong version on two accounts: it asserts a non-defeasible connection of conceptual necessity between moral judgements and endorsement and a non-defeasible connection of conceptual necessity between moral judgements and background conative-affective dispositions. There is a strong sense, then, in which endorsing attitudes and conative dispositions are 'internal' to moral judgements. The full account of this version of internalism developed in chapters 3 to 5 is more complicated than the simple version of internalism with which I started this section. In order to understand the need for this more complex version of moral judgement internalism we will need start with the simple version and work our way back to the more sophisticated version. In chapters 2 and 3, I assume something like the simple version of moral judgement internalism in order to examine some of the central issues dividing internalists and externalists. My approach is to establish cognitivist-internalism as a *prima facie* plausible position consistent with the desire/belief thesis. This is intended to address standard objections to cognitivist-internalism that dismiss it out of hand because it does not accommodate the desire/belief theory of action explanation and motivation.

The application of speech act theory to the problem of moral motivation in chapter 3 shows that it is *prima facie* plausible that moral judgements express both beliefs and motivating attitudes. Such motivating attitudes can take part in standard desire/belief accounts of action explanation. However, I argue in chapter 4 that we need a more developed account of the role and nature of the motivating attitudes involved in moral judgement and deliberation, for the reasons outlined at the start of this section. In chapter 5, I call into question the adequacy of the standard desire/belief model in dealing with the complex structure of moral motivation outlined in chapter 4. In attempting to accommodate the simple model we come to a better understanding of

the most plausible way of interpreting the internalist thesis and this leads to a revision of the simple model.

## **7. Conclusion**

Is it possible to address the central issues raised in this chapter without engaging fully in the debate between those who think that only desire can motivate and those who hold a cognitivist theory of motivation? Ultimately, I think the answer has to be no. However, what we can do is move back from the debate about moral psychology and approach it from a different direction in the hope of understanding it more clearly. I attempt this in chapters three and four by applying the theory of speech acts to the problem of moral motivation with the aim of developing a workable union of cognitivism and moral judgement internalism. Before doing this, however, we need to be convinced that the difficult task of combining cognitivism and internalism is a necessary one. Perhaps the intuitions which appear to suggest internalism can be accommodated by an externalist account of moral judgement, thereby removing the need to develop the speech act model I propose or a revisionary theory of motivation in place of the desire/belief model. It is to an examination of moral judgement externalism that I now turn.

---

## Externalism and Moral Motivation

### 1. Introduction: Moral Judgement Externalism

Externalists about moral motivation reject the internalist claim that there is a connection of conceptual necessity between moral judgements and motivation and assert the possibility of making genuine moral judgements in the absence of motivation.<sup>1</sup> They argue that moral motivation is contingent upon an agent's having an independent desire to perform the acts she judges to be right. Most defenders of externalism attempt to discredit the claim of internalism to reflect a common intuition about morality by formulating a counterexample to the internalist thesis. This usually involves a case where we feel confident in ascribing a moral judgement to an agent even though she fails to be motivated by it. This possibility is said to be typified by the character of the amoralist: 'someone who recognises the existence of moral considerations and remains unmoved' (Brink 1989: 46).

It is commonly held that moral judgement externalism is simply the denial of moral judgement internalism (e.g. Brink 1989: 42).<sup>2</sup> However, as it has developed in the literature externalism involves more than this denial. I noted in chapter one that Nagel rejects both the claim that moral judgements entail motivation and the desire/belief model of motivation favoured by externalists. Nagel is neither a moral judgement internalist nor a moral judgement externalist, at least in the sense in which the latter is understood in the literature. Externalists do not merely deny that moral judgements entail motivation, they deny that moral judgements are capable of motivating *at all* in the absence of a prior and independent desire of the relevant kind (Brink 1989: 42; Railton 1986: 168-170; Svavarsdóttir 1999: 170). Nagel, by contrast,

---

<sup>1</sup> For example: Boyd 1988; Brink 1989: chap. 3, 1997; Copp 1997; Railton 1986; Svavarsdóttir 1999; Dreier 2000; Shafer-Landau 2000.

<sup>2</sup> In general, I will abbreviate 'moral judgement internalism' and 'moral judgment externalism' to 'internalism' and 'externalism' respectively, and will refer to other forms of internalism by their full titles. Occasionally, 'internalism' and 'externalism' will refer to a more general orientation towards one side or the other which may encompass a number of different internalist or externalist theses in ethics. This use will be made clear by the context in which it occurs.

argues that, while moral judgements do not entail motivation, such judgements are capable of motivating without the presence of a prior and independent desire. Externalism is thus best understood as an attempt to combine cognitivism with the desire/belief model of motivation. It does this, and provides its solution to the problem of moral motivation, by rejecting moral judgement internalism. This allows externalists to accept that no fact or cognitive state is capable by itself of entailing (or producing) motivation without thereby accepting that this in any way impugns the cognitive nature of moral judgements.

Much of the externalist resistance to internalism derives from the threat internalism is thought to pose to cognitivism, and by extension to naturalism and realism in ethics (e.g., Brink 1989: 37; Shafer-Landau 2000: 270). Externalists generally accept the validity of the non-cognitivist anti-cognitivist argument and so reject the internalist premise in that argument to defend cognitivism (Brink 1989: 44-45; Svavarsdóttir 1999: 168-169). In chapter one, I argued that the anti-cognitivist argument is invalid and that internalism does not pose the kind of threat to cognitivism this argument suggests. The aspects of morality externalists seem most concerned to accommodate can be made consistent with what I have termed act internalism. For example, in addition to the threat posed by non-cognitivism, externalists worry that a connection of conceptual necessity between moral judgements and motivation might threaten the categoricity of moral demands by supporting some form of subjectivism, or what I have called content internalism about motivation (cf. Brink 1989: 45; Railton 1986: 166-171). This would occur if the truth of moral claims were dependent on the desires of those who made them. However, because act internalism is a thesis about the act of making moral judgements and not about the content of moral judgements it in no way threatens the independence of moral truth from contingent desires, and thus in no way threatens the categoricity of moral demands.

Much of what externalists want to capture about morality can be accommodated within the framework of some form of act internalism. This is fortunate, because externalism fails to account for the commendatory aspect of moral judgements and,

more generally, the interested or participatory relation between agents and their moral deliberations. I detail these criticisms of externalism in section five below. Section two begins with an outline of the two standard externalist models of moral motivation. One of these explains moral motivation in terms of desires with non-moral content, and the other explains such motivation in terms of desires with moral content. I argue that the non-moral content model fails to account for the way in which motivation tracks moral judgements.

In section three I discuss the externalist argument from amorality. The amoralist is a figure who is said to make genuine moral judgements yet remain unmoved. Externalists argue that the possibility of such a figure undermines internalism's conceptual claim. I stress the importance of distinguishing between amorality as a thesis about motivation and amorality as a thesis about justification. In addition, I argue that it is crucial to identify the precise concept involved in the internalist's conceptual claim. Section four involves a lengthy discussion of one of the central debates in the literature on moral motivation, between Michael Smith and David Brink. Smith presents a range of objections to externalism and to the amoralist argument.<sup>3</sup> One of his central claims is that externalism cannot account for the reliable connection between moral judgements and motivation without invoking a model of motivation that gives a false account of the motivational psychology of virtuous agents. I argue against Smith that his objection threatens to make genuine moral motivation impossible, and, moreover, that it applies equally to his own model of moral motivation.

As already mentioned, section five raises a number of objections against externalism. I argue that externalism provides an unconvincing explanation of the practical or action-guiding aspect of moral judgements because it fails to account for the participatory nature of moral deliberation. Focussing on the point or function of moral deliberation enables us to see why someone outside morality, such as the amoralist, is unable to make genuine moral judgements. I conclude that we have

---

<sup>3</sup> See Smith 1994: esp. chap. 3, and 1997.

reason to prefer some form of act internalism to either externalism or to the kind of content internalism espoused by Smith.

## **2. Externalist Models of Moral Motivation**

There are two main models of externalist moral motivation: one explains moral motivation in terms of a desire or desires with non-moral content, and the other appeals to a desire with specifically moral content to explain such motivation. Both varieties are externalist, holding that the motivating desires in question are only contingently and non-conceptually related to agents' moral judgements. I argue in this section that the non-moral content model fails to explain effectively how motivation tracks moral judgements.

Brink advances a version of the non-moral content model of moral motivation. He argues that the connection between moral judgements and motivation is dependent both on the content of the moral views to which an agent subscribes and to the content of the desires the agent happens to have:

[E]xternalism can base [moral] motivation on “deep” or widely shared psychological facts. Let’s think about common moral views that recognise the other-regarding character of many moral demands. If, for example, sympathy is, as Hume held, a deeply seated and widely shared psychological trait, then, as a matter of contingent...psychological fact, the vast majority of people will have at least some desire to comply with what they perceive to be their moral obligations, even with those other-regarding moral obligations. (Brink 1989: 49)<sup>4</sup>

The problem with this model is that it fails to explain the way in which motivation reliably tracks moral judgements, and this is something that must be accounted for by any adequate theory of moral judgement (cf. Smith 1994: 71). Given the complex nature of moral requirements, it is unlikely that merely having sympathy towards others will account for the way in which changes in motivation track changes in moral judgement. For example, suppose that John believes it is morally wrong to

---

<sup>4</sup> Boyd (1988: 215) defends an externalist account of the role of sympathy. He argues that it is possible that sympathy plays a cognitive as well as a motivational role by giving agents the capacity to recognize human goods and harms to which they might otherwise be insensitive.



hold asylum seekers in detention centres, and that he is motivated to break the law to help those held in detention out of a deep concern for their well-being. John is a kind person with sympathy and concern for the suffering of others. However, after numerous long conversations with a friend who works for the government, John comes to believe that the policy of enforced detention of asylum seekers is morally right. Let's say that the arguments which persuade him are highly abstract and don't invoke his sympathy in any way. Nevertheless, we would expect John's motivations to track his moral judgement and for him to be motivated to stop undermining the system of enforced detention. John still feels concern for the well-being of the detainees, but now believes that he is morally obliged to distance himself from their plight. If John's motivations do change to reflect his changed moral belief, we can no longer appeal to his sympathetic nature to explain why he is motivated to do what he believes is right. Something else will need to be appealed to explain his new motivation.

Shafer-Landau suggests that the externalist need not be restricted to one kind of desire to explain the reliability of moral motivation (2000). He argues that the externalist can appeal to a range of non-moral desires to explain why someone is motivated by her moral judgements. Judgements of moral obligation generally will require an agent either to perform an act that furthers her own interests or the interests of others. If the required act furthers the agent's own interests we can appeal to the agent's 'standing self-concern' to explain why she is motivated to perform it. If the required act conflicts with an agent's self-interest, we can appeal to a range of other-regarding desires – to be kind, to avoid harming others, and so on – 'to explain why our motivations track our moral judgements' (Shafer-Landau 2000: 285.).

The difficulty with Shafer-Landau's suggestion is that, contrary to his last statement, it provides no convincing explanation of how motivations *track* moral judgements. On Shafer-Landau's model, if I judge, for example, that I am morally required to take my library books back on time, we can explain why I do so by appealing to my self-interested desire not to be fined. But suppose that the library makes me exempt

from fines because of some kind of administrative mistake, so that I can now bring books back late with impunity. Why expect that I will still be motivated to take the books back? According to Shafer-Landau, it is likely that I will have other-regarding desires that will be able to explain why I continue to take my books back – that can take over from my self-interested desires. Perhaps I decide that I don't want to deprive others of the books. But notice that the fact that I think that taking the library books back is *morally right* is irrelevant to my decision making. To see this we simply have to imagine what happens if I decide that I have no moral obligation to take my library books back. If I don't want to be fined, then I will still desire to take the books back on time, and I would desire to take the books back on time even if exempted from fines, if I had a desire not to deprive others of books. So my motivations do not track my moral judgements; they track my interests, which may or may not coincide with the acts I judge to be right. I take the inability of non-moral content versions of externalism to explain the way in which motivation tracks moral judgements to undermine such approaches.

Externalists can avoid such difficulties and explain the way in which motivation tracks moral judgements by invoking a desire with specifically moral content, such as the desire to do the right thing, or the desire to be moral (e.g., Svavarsdóttir 1999: 170; cf. Shafer-Landau: 286). Suppose that John has a standing desire to do what is morally right. If he judges that it is morally right to break the law to help asylum seekers, then we would expect him to be motivated to do so. And we would have this expectation regardless of whether he normally shows a great deal of sympathy towards others. When John's moral belief changes after talking to his friend, and he comes to believe that he is morally required to desist from helping the asylum seekers, we will, given his standing desire, expect him to be motivated to stop helping the detainees. So the moral content version of externalism provides a simple explanation of why motivation reliably tracks moral judgements while upholding the view that the connection between moral judgements and motivation is an entirely contingent matter. In section four, I discuss Smith's fetishism objection to this kind of externalism. He argues that while moral content versions of externalism can explain the tracking condition on moral motivation, in doing so they give a false

account of the motivational psychology of virtuous people. I now turn to one of the central arguments advanced in favour of externalism.

### 3. The Externalist Argument from Amoralism

I said above that externalists often attempt to discredit internalism by formulating a case where we feel confident in ascribing a moral judgement to an agent even though she fails to be motivated by it.<sup>5</sup> This is the approach adopted by Brink in his defence of moral judgement externalism. He argues that internalism is implausible because it makes this kind of failure of motivation conceptually impossible:

[I]nternalism, so construed, seems just false to both actual and possible psychological facts. Although indifference to what is regarded as moral considerations may be fairly rare, it does seem to exist. Some people (e.g., certain sociopaths) do not care about what they regard as moral consideration...The internalist about motives [moral judgement internalist] claims it is a conceptual truth about morality that moral judgement or belief motivates. According to the internalist, then, it must be conceptually impossible for someone to recognise a moral consideration or assert a moral judgement and remain unmoved. This fact raises a problem for internalism; internalism makes the amoralist conceptually impossible. (Brink 1989: 46)

Brink's definition of the amoralist as someone who is unmoved by her moral judgements is largely stipulative, although it seems reasonably close to the standard definition of an amoralist as someone who is unconcerned whether something is right or wrong.<sup>6</sup> The definition appears to cover the group outlined by Michael Stocker, who fail to be motivated by their moral judgements as a result of depression, fatigue or weakness of will (1979). Stocker argues that the mood, interest, and energy of an agent can have profound effects on her ability to be motivated by her evaluative judgements. He holds that the mediating role played by these moods and interests confutes the claim that what is judged good must necessarily attract. The idea is that when we suffer from depressions, tiredness, resentment, despair, and the like, we often fail to be motivated to do what we genuinely judge to be good. Brink tends to concentrate on more extreme examples of amoralism, such as the sociopath:

---

<sup>5</sup> Svavarsdóttir's (1999) example of 'Patrick' takes this form, as do a number of examples in Shafer-Landau (2000: 272-274).

<sup>6</sup> The Concise Oxford Dictionary (Tenth edition) gives the following definition for 'amoral': "lacking a moral sense; unconcerned whether something is right or wrong."

individuals with a general disregard or contempt for morality. It is not immediately clear why Brink opts to defend the possibility of such an extreme figure when all that is required to undermine internalism is the conceptual possibility that an agent is on one occasion too depressed to be moved by something she genuinely takes herself to be morally required to do.

One reason I think Brink focuses on extreme examples of amorality is that much of his discussion is directed against content internalism about justifying reasons rather than merely against moral judgement internalism. In order to undermine the former kind of internalism Brink introduces the notion of ‘amoralist scepticism’. This involves someone who ‘accepts the existence of moral facts and concedes that we have moral knowledge, and asks why we should care about these facts’ (1989: 46).<sup>7</sup> Brink views the amoralist sceptic as someone who questions the *justification* or *rationality* of moral demands (Brink 1989: 48). Stocker’s depressed agent is not someone who necessarily calls into question the justification of moral demands. She may readily grant that she has a, perhaps overriding, justifying reason to conform to the moral demand in question, but she simply lacks the energy to be moved to do so. Such an agent is not sceptical about the justification or authority of moral demands and is thus perfectly consistent with content internalism about justifying reasons.

Brink wants to deny both content internalism about justifying reasons and moral judgement internalism and suggests we should accept the possibility of motivational failures not linked to issues of justification, what he calls ‘unprincipled amorality’, and amoralist scepticism.<sup>8</sup> But the possibility of amoralist scepticism doesn’t by itself undermine moral judgement internalism: it might be conceptually necessary that in making a moral judgement an agent is motivated to conform to it, yet still possible for the agent to question whether such motivation is justified (cf. Brink 1989: 48). So amoralist scepticism is neither sufficient nor necessary to undermine

---

<sup>7</sup> Part of the difficulty here is that the notion of being ‘unconcerned’ about moral considerations is ambiguous. I might not care about something because I see no reason to care, or I might not care about something because I simply don’t have the energy to care, even though I believe I have reason to do so.

<sup>8</sup> Brink means by the unprincipled amoralist someone who fails to be motivated but does not question the rational justification of her moral judgement.

moral judgement internalism. It is important that we keep the issue of whether moral judgements are conceptually tied to justifying reasons separate from the issue of whether moral judgements are conceptually tied to motivation. The former issue is only really plausible as a thesis about the content of moral judgements, whereas the latter issue can be interpreted as a thesis about the act of making a moral judgement as well as about the content of moral judgements.<sup>9</sup>

A standard way for the internalist to respond to purported examples of motivational failure is to argue that while it may appear that the unmoved agent is making a moral judgement she is in fact doing no such thing (Hare 1952: 124; Smith 1994: 67; cf. Brink 1989: 46). The most well known argument along these lines is Hare's inverted commas argument. According to Hare, what distinguishes moral judgements from purely descriptive judgements is their commendatory function. In so far as a judgement that something is right or good is a genuine evaluative judgement it must express the agent's motivation to do or promote the favoured option. Brink's amoralist while giving the impression of making an evaluative judgement is doing no such thing because she fails to commend or be motivated by the option she judges right. At most she is describing what other people take to be good or right as 'good' or 'right' (Hare 1952: 124). Hare calls this the inverted commas use of evaluative terms. It occurs when we allude to the evaluations of others without commending those evaluations.

Brink argues that the inverted commas response fails to take the amoralist challenge seriously enough (1989: 47). He says that we cannot simply dismiss the possibility of someone who claims to believe that an action is right but who fails to be moved by her judgement, as such a person is a familiar character historically in discussions of such matters.<sup>10</sup> Svavarsdóttir gives a more detailed example of such an agent. She suggests that we should think about the example in terms of which hypotheses might plausibly explain the observable behaviour described:

---

<sup>9</sup> Although I argued in chap 1 that we should interpret moral judgement internalism as a thesis about judgement acts rather than judgements contents, the latter is at least *prima facie* plausible, particularly if we focus on the notion of a *conceptual* connection rather than that of a necessary connection. By contrast, it is difficult to see how the issue of justifying reasons could fit the judgement act model.

<sup>10</sup> Brink mentions Plato's Thrasymachus and Hobbes's Fool as examples of amoralists (1989: 47).



Virginia has put her social position at risk to help a politically persecuted stranger because she thinks that it is the right thing to do. Later she meets Patrick, who could, without any apparent risk to himself, similarly help a politically persecuted stranger, but who has made no attempt to do so. Our morally committed heroine confronts Patrick, appealing first to his compassion for the victims. Patrick rather wearily tells her that he has no inclination to concern himself with the plight of strangers. Virginia then appeals to explicit moral considerations: in this case, helping strangers is his obligation and a matter of fighting enormous injustice. Patrick readily declares that he agrees with her moral assessment, but nevertheless cannot be bothered to help...Later he shows absolutely no sign of regret for either his remarks or his failure to help. (Svavarsdóttir 1999: 176-177)

Svavarsdóttir argues that the burden of argument is on those who want to restrict the range of hypotheses in the running as explanations of such examples. She contends that this is precisely what internalism as an a priori constraint on accounts of moral judgement does and that, therefore, the burden is on internalism to justify such a restriction (Svavarsdóttir 2001: 181).

The committed internalist will generally insist that externalist interpretations of such examples are question-begging. As we will see in section four, this is precisely Smith's response to Brink's argument (1994: 68-71). Svavarsdóttir appears to avoid this charge as she advances externalism merely as one possible explanation of Patrick's behaviour. She does take externalism to be the most plausible explanation of his behaviour but says she is more interested in the notion that it cannot be readily ruled out as false or defective (Svavarsdóttir 1999: 179). Nonetheless, she does assume that in order for Patrick to be sincere in making his moral judgement he only needs to believe the relevant moral proposition. What is interesting about Patrick is that he never gives moral considerations as reasons for his actions, and he engages in what he admits is morally wrong behaviour without displaying signs of hesitation, regret, guilt or shame (Svavarsdóttir 2001: 178). Patrick does not aspire to live by moral standards and tends to feel regret and shame when he fails in relation to things he finds important rather than in relation to moral considerations (Svavarsdóttir 2001: 178).

While we may agree with Svavarsdóttir that it is plausible to ascribe moral beliefs to Patrick, it is not clear why we should accept that the possession of such a belief in



the absence of any connection to moral emotions, any connection to a practical function, and any acceptance of even minimal normative authority should count as a fully-fledged moral judgement. Svavarsdóttir touches on this point when she notes that it ‘seems reasonable to operate with the assumption that an individual who bothers to make moral judgements is morally committed to some degree’ (2001: 185). She then concedes that motivation falls out of the concept of commitment. The issue then becomes whether our standard concept of a sincere moral judgement includes some notion of commitment.<sup>11</sup> We are back with a conceptual claim; but one that cannot be decided merely by noting that it is plausible that Patrick possesses a moral belief.<sup>12</sup>

At least one theorist has described the debate between internalists and externalists as ‘at a standoff’ of a sort that is ‘depressing’ in philosophy (Dreier 2000: 620). I think this is not an uncommon response to a debate in which both sides take their intuitions as bedrock, without, perhaps, exploring the possible sources of their disagreement. At the centre of this disagreement, as we have seen, is a conceptual thesis. It states that the necessary connection between moral judgements and motivation posited by the internalist derives from the *concept* of moral judgement: it is part of what we mean when we say that someone has made a moral judgement that she is motivated to act in accordance with it. It is because internalism is advanced as a conceptual thesis that arguments for and against it rely heavily on appeals to intuition about examples in which an agent claims something to be morally right or wrong.

Given that the internalist’s conceptual thesis deals with the issue of whether a specific concept has a motivational component as part of its meaning *it would seem*

---

<sup>11</sup> In chapters 3-5, I argue that if an agent makes a sincere moral judgement then she commits herself to being guided by the requirement expressed by the judgement.

<sup>12</sup> David Copp argues that there is a complex state of moral conviction that entails motivation but that motivation is external to moral belief (1995: 205-206). In this 1995 article Copp defends a form of moral judgement externalism and, like Svavarsdóttir, seems to think that this position is vindicated by externalism about moral belief and motivation. In chapter 4, I discuss a more recent article by Copp which is more sympathetic to moral judgement internalism.

*to be crucial to determine from the outset the precise concept involved.* There seem to be two main candidates for the item that moral judgement internalism states is necessarily tied to motivation. One candidate is the concept of making a moral judgement and the other the concept of the content of moral judgement. This is the distinction between judgement act and judgement content outlined in chapter one. I gave a range of reasons there for interpreting moral judgement internalism as a thesis about judgements acts rather than judgement contents.<sup>13</sup> However, in contemporary discussions this distinction is not commonly noted.<sup>14</sup> This means that it is often not clear what the precise concept is about which internalists and externalists disagree. The issue is further complicated by the fact that internalism as a doctrine about moral judgements and motivation and internalism as a doctrine about moral judgements and justifying reasons are often not kept apart, the intuitions that support the latter often being used to support the former (e.g., Smith 1994: 6).

In the next section, I look at how these issues play out in one of the central debates in the literature on moral motivation in an attempt to come to a better understanding of the precise nature of the disagreement between internalists and externalists. It will become clear that act internalism and content internalism have very different implications for the explanatory adequacy of internalist theories of moral motivation and thus different levels of success in responding to the challenge of externalism.

#### **4. Smith's Anti-Externalism**

##### *i) The analogy with colour judgements*

Smith has reservations about Hare's inverted commas argument because he thinks it unlikely that those who fail to be motivated by their moral judgements necessarily describe what other people take to be right and wrong. However, he thinks Hare's approach is substantially correct: 'The point is not that amoralists really make judgements...about what other people judge to be right and wrong...The point is

---

<sup>13</sup> For example, act internalism is platitudinous but content internalism is not, content internalism threatens the category of moral demands, and act internalism but not content internalism captures the commendatory function of moral judgements.

<sup>14</sup> Svavarsdóttir mentions the distinction briefly: 'On [one] view, motivational internalism [moral judgement internalism] falls out of the concept of a moral judgement rather than out of the concepts employed in moral judgements' (1999: 182).

rather that the very best we can say about amoralists is that they try to make moral judgements but fail' (Smith 1994: 68). Smith attempts to show why this claim is not '*ad hoc*' by developing an analogy with colour judgements. It is not clear how the colour analogy is meant to support the inverted commas interpretation of the amoralist argument. Its main point seems to be to show that Brink's use of the example of amoralism to support externalism is question begging (Smith 1994: 70).

Smith argues that it seems intuitively plausible that the possession of colour concepts and mastery of colour terms depends on an agent's having the appropriate visual experiences. Smith gives the example of a congenitally blind person who is able to engage reliably in colour ascription through non-visual awareness of surface reflectance properties, and who is thus able to use colour terms reliably. The thought is that we would be reluctant to say of this person that they were making genuine colour judgements, though we might say that they were making 'colour judgements'. The blind person who reliably uses colour terms is meant to be analogous to the amoralist. The amoralist, Smith concedes, may be able to use moral terms reliably to refer to the same descriptive properties referred to by people making genuine moral claims; but just as the blind person fails to make genuine colour judgements because she fails to have full mastery of colour terms, so the amoralist fails to make genuine moral judgements because she fails to have full mastery of moral terms:

[Internalists say] that a subject has mastery of colour terms (moral terms), and thus really makes colour judgements (moral judgements), only if, under certain conditions, being in the psychological state that we express when we make colour judgements (moral judgements) entails having an appropriate visual experience (motivation). (Smith 1994: 70)

Smith charges the externalist with question begging: in advancing the amoralist as a counterexample to internalism, the externalist merely assumes that mastery of moral terms does not include being motivated (1994: 70).

Before addressing Smith's use of the colour analogy, I want to comment on his debate with Brink and how it relates to his use of Hare's inverted commas argument. The debate between Brink and Smith about the possibility of amoralism is a central and influential debate in the literature on moral motivation and internalism.

However, assessing the debate is made difficult because Brink and Smith actually work with quite different accounts of the connection between moral judgements and motivation. Smith's 'practicality requirement' states that if an agent judges that something is morally right, then either she is motivated to do it or she is practically irrational (1994: 61). By contrast, Brink's version of internalism states that moral judgements entail motivation (1997: 7-8). Smith adopts his formulation because he accepts the argument that agents can fail to be motivated by their moral judgements when sufficiently depressed, tired, or weak-willed. But this means that on Smith's model it is perfectly possible to make a sincere and genuine moral judgement without being at all motivated to act on it; and this is the denial of what I have called moral judgement internalism. So Smith and Brink are in agreement about the possibility of unprincipled amorality.

It is easy to forget that Smith and Brink both accept unprincipled amorality as they take up opposing sides with regard to the traditional debate between non-cognitivist internalism and externalism. Strictly speaking, Smith should side with Brink against Hare: On Smith's theory it is possible to make a genuine moral judgement without commending, prescribing, or expressing a motivating state of any kind. Smith accepts the denial of moral judgement internalism. Difficulties arise because Smith presents his practicality requirement as a weaker version of moral judgement internalism, and this characterisation is generally accepted by his opponents.<sup>15</sup> But it is central to Smith's conceptual claim about motivation that internalism involves the *content* of moral judgements (cf. Copp 1997: 37). This is why Smith says that someone who fails to be motivated by her moral claim does not have mastery of the relevant moral terms (1994: 70). Smith's developed theory is an attempt to secure a conceptual connection between moral judgements and motivation by providing a substantial theory of the content of such judgements.<sup>16</sup>

---

<sup>15</sup> For example, Svavarsdóttir accepts the weaker thesis, although she thinks we should not characterise such failures of motivation in terms of failures of rationality, but in terms of motivational disorders more generally (1999: 164). However, she still seems to regard Smith's practicality requirement as a 'close relative' of moral judgement internalism.

<sup>16</sup> He defends 'rationalism' or what I have termed content internalism about justifying reasons (Smith 1994: 85-91). This is the view that it is part of the concept of a moral requirement that such requirements provide justifying reasons for action. Smith combines this with an account of what it is to have a justifying reason for action: an agent has a justifying reason to perform an act if her fully

In chapter one I argued that any theory which attempts to secure the connection between moral judgement and motivation at the level of content is a version of content internalism rather than moral judgement internalism. And this is precisely what Smith does with the practicality requirement. But this means that the difference between the practicality requirement and non-cognitivist versions of moral judgement internalism, or what I have termed act internalism, is a difference in kind rather than degree: the former is about judgement contents and the latter about judgement acts.<sup>17</sup> This is particularly confusing in relation to Smith's debate with Brink because it is clear that Brink views moral judgment internalism as distinct from content internalism about motivation (1989: 40-42).<sup>18</sup>

If it is true that making a genuine colour judgement entails having the appropriate visual experience then Smith's analogy fails. This is because, as we have seen, Smith denies that moral judgements entail motivation. To be structurally analogous, colour judgements would have to be connected defeasibly to the appropriate visual experience; that is, it would have to be possible to make genuine colour judgements without having the appropriate visual experience. Furthermore, this condition on colour judgements would need to derive from the content of those judgements. One way this might work is if he adopted a response-dependent account of colour

---

rational self would desire her to perform the act (1994: chap 5). On this account, if an agent judges that she is morally required to perform a certain act, then she judges that if fully rational she would desire herself to perform the act. Consequently, if the agent fails to be motivated to perform the act, she is, says Smith, practically irrational (1994: 62, 148).

<sup>17</sup> The issue is further complicated by the fact that the practicality requirement is only a version of content internalism about *motivation* because Smith's theory of justifying reasons means that motivation is included in the content of reason judgements (see note 13). A content externalist about reason judgements and motivation who accepts the practicality requirement will deny that moral judgements (as reason judgements) have an internal connection to motivation. Consequently, the practicality requirement is consistent with a thorough-going externalism about moral judgements and motivation. That is, the practicality requirement is not, as it stands, necessarily a version of internalism about motivation at all, weak or strong.

<sup>18</sup> Brink's 'appraiser internalism' about motivation is equivalent to what I call moral judgment internalism and his 'agent internalism' about motivation is equivalent to my content internalism about motivation (1989: 40). Brink argues, as I did in chap 1, that content internalism about motivation is far less intuitively plausible than moral judgement internalism (1989: 41, 45). This is because it makes moral requirements dependent on whether or not agents happen to desire to conform to them; and this goes against the widely held view that moral requirements are categorical, in the sense that they apply to agents independently of their contingent desires. It is therefore highly implausible that the internalist intuition is about this kind of content internalism.



concepts.<sup>19</sup> For example, for something to be red might be for it to look red to standard perceivers under normal conditions. But if this were the case then it should be possible for the blind person to make colour judgements, as long as part of what she judged was that the object in question would look red to standard perceivers under normal conditions. Now, we might worry that a congenitally blind person would not be able to form a concept of what it is for an object to look coloured. Even if we grant this, however, it still leaves room for those who become blind or who are temporarily blind to make genuine colour judgements (cf. Brink 1997: 24).

If we adopt the response-dependent model of colour concepts, we should accept the possibility of someone who is blind – if only temporarily or not congenitally – making genuine colour judgements. This is supported by the fact that there seem to be numerous everyday examples of people making colour judgements without having *at the time of making the judgement* the appropriate visual experience. If I tell you that the walls in my bathroom are green, I don't need to be looking at my walls for this to count as a genuine colour judgement. Of course, the judgement is parasitic on my having had the appropriate visual experience, or at least normally would be parasitic. But imagine that my partner renovates the bathroom while I am overseas and rings me up to tell me that she has painted the bathroom green. Later in the day I tell you about the renovations and mention that the bathroom walls are green. Now, when I make this judgement I am not having, and have not had, the visual experience of seeing the colour of my bathroom walls, yet it is typical, surely, of genuine colour judgements that people make everyday. Adopting the response-dependent account of colour concepts is one way of making the colour judgement example structurally analogous to Smith's account of moral judgements. But it involves a particular account of the content of colour judgements, just as Smith's practicality requirement involves a particular account of the content of moral judgements. It is not by itself an argument for that account of the content of colour judgements, and has no direct bearing on the adequacy of Smith's account of the content of moral judgements.

---

<sup>19</sup> For a discussion of this idea in relation to response dependent accounts of value see McDowell 1985 and Wright 1988.



Although Smith claims that the amoralist argument is question-begging, it is not clear that the externalist move is as crude as he suggests. The amoralist example is meant to be compelling intuitively regardless of one's prior convictions about the relationship between moral judgement and motivation (cf. Brink 1997: 22-23). The externalist holds that if the force of this intuition is greater than that of the internalist intuition, we have good reason to reject internalism about moral judgement. The problem is not that the amoralist move is question-begging *per se*, but that there seem to be few theoretically neutral intuitions we can rely on to interpret putative examples of amoralism.<sup>20</sup> This merely reinforces the importance of being clear about the particular conceptual connection at issue. What Smith's colour judgement analogy does make clear is that he views internalism as a thesis about the content of moral judgements. Importantly, Smith does recognise that such examples are inconclusive and the need to move beyond them. This leads him to advance an independent positive argument against externalism.

*ii) The Reliability Thesis*

Smith's positive argument against externalism begins with the observation that changes in motivation reliably track changes in moral judgement:<sup>21</sup>

By all accounts, it is a striking fact about moral motivation that a change in motivation follows reliably in the wake of a change in moral judgement, at least in the good and strong-willed person. A plausible theory of moral judgement must therefore explain this striking fact. As I see it, those who accept the practicality requirement can, whereas strong externalists cannot, explain this striking fact in a plausible way. (Smith 1994: 71)

He uses an example of a change in moral motivation to elucidate his charge against externalism. Suppose that I am engaged in a political debate with someone prior to an election. I have always voted for the libertarians but my interlocutor persuades me that the values that the libertarians promote, and which I have always adhered to, are fundamentally mistaken. They convince me to alter my moral beliefs and to vote for the social democratic party. This change in moral beliefs is followed reliably by a change in motivation such that I am now motivated to vote for the social

---

<sup>20</sup> For example, see my discussion of Svavarsdóttir's 'Patrick' example above.

<sup>21</sup> I will refer to this claim as the 'reliability thesis'.

democratic party. Smith thinks that there are only two options which adequately explain this reliable connection:

On the one hand we can say that the reliable connection between judgement and motivation is to be explained internally: it follows directly from the content of moral judgement itself...Or, on the other hand, we can say that the reliable connection between judgement and motivation is to be explained externally: it follows from the content of the motivational dispositions possessed by the good and strong-willed person. (Smith 1994: 72)

*Prima facie* it seems that the externalist has a straightforward and unmysterious explanation of the reliability of moral motivation in the 'good and strong-willed person'. The good person is plausibly someone who has a standing desire to do what she takes herself to be morally required to do. And the strong-willed person is someone in whom this desire is reliably effective in motivating her to do what she judges she is required to do. So the externalist model conforms easily to the desire/belief account of motivation. In the above example, my motivation changes from the libertarians to the social democrats because I change my belief about which party advances morally correct policies. Given that I have a standing desire to do what is morally required, I now believe that this desire will be satisfied by my voting for the social democrats rather than the libertarians. And given that I am strong-willed, I will now be motivated to vote in the way I judge will satisfy my standing desire to do what is morally required. Structurally, this explanation of motivational change is the same as the basic desire/belief model's explanation of how changes in instrumental beliefs lead to changes in motivation.

According to Smith there are a number of disturbing implications that follow from the externalist explanation of the reliability of moral motivation. He argues that if moral agents were motivated in the way suggested by externalists, then such motivation would be akin to a fetish or moral vice rather than a moral virtue (Smith 1994: 71-76). The fetishism argument is meant both to undermine externalism and to vindicate the practicality requirement. However, even if the argument against externalism is successful, it is not clear that it directly supports the practicality requirement rather than some other account of the reliability thesis. Smith seems convinced that externalism and the practicality requirement are *the only two* options

capable of explaining the reliability of moral motivation (1994: 72, 76). But we need not accept this claim. The reliability might be explained by something other than the content of moral judgement, as given by the practicality requirement, or the content of an external desire. Specifically, it might be explained in relation to the conditions governing the performance or act of making a moral judgement. In fact, non-cognitivist internalism, as a type of act internalism, already provides us with an alternative to Smith's choice.

The central claim of Smith's argument is that internalism provides a significantly more plausible explanation of the reliability of moral motivation than does externalism. But some internalist explanations of the reliability thesis are more compelling than others. Smith notes that internalist explanations can be either cognitivist or non-cognitivist (1994: 72). On the cognitivist alternative, an agent's moral motivation derives from her belief that something is good or right: the belief somehow brings about the appropriate motivation. On the non-cognitivist alternative, moral judgements express desires or some other kind of pro-attitude rather than beliefs and so are essentially motivating. Obviously, choosing the cognitivist rather than the non-cognitivist version of internalism, or vice versa, will have implications for the kind of explanation the practicality requirement provides for the posited connection between moral judgement and motivation.

Smith seems to think that the non-cognitivist alternative is an example of the reliability of moral motivation being explained by the content of moral judgement and that its success vindicates the practicality requirement (1994: 72). But non-cognitivism tells us that having the appropriate motivation is constitutive of the *act* of making a moral judgement: expressing a pro-attitude or commending is part of what we are *doing* when we make a moral judgement; it is not part of the content of the moral judgement. I think that non-cognitivism, as a form of act internalism, can claim to provide a *prima facie* plausible account of moral motivation. Much of the appeal of the non-cognitivist version of internalism lies in its ability to explain a variety of the features common to ordinary moral practice by appealing to the widely accepted desire/belief model of action explanation. But non-cognitivism does not

support the practicality requirement. If an agent fails to endorse or commend an option she judges right and thus fails to be motivated to do it, non-cognitivism tells us that she has not made a genuine moral judgement, not that she is practically irrational.

The point to stress is that we can only assess the explanatory adequacy of specific examples of internalism. It is not enough to say that an explanation of the reliability of moral motivation follows as a 'direct consequence' of the truth of internalism or the practicality requirement (Smith 1994: 72). When Smith criticises externalism, he criticises a specific account of how moral judgements motivate. Internalism offers no real explanation if all it offers is a statement of the internalist thesis without an account of how on such a model moral motivation actually works. If, as I believe, act internalism provides a distinct alternative to Smith's version of internalism, then the argument based on the reliability thesis fails to vindicate the practicality requirement and fails to deal with a significant rival.

*iii) Smith's Model of Moral Motivation*

How exactly on Smith's 'rationalist' model do moral judgements motivate? He argues that they motivate by virtue of their content, so that 'the belief that an act is right produces a corresponding motivation' (Smith 1994: 72). Smith's account of moral motivation is initially surprising given his adherence to the desire/belief or 'Humean' model of motivation, as it implies that belief rather than desire can be the principal source of motivation. This tension in Smith's account is not often recognised. For example, while Brink is broadly opposed to Smith's internalism, he does not seem to consider Smith's account of moral motivation to be in conflict with the desire/belief model in the way, for example, that Nagel's version of rationalist motivation is in conflict with it (1997: 15).

According to Brink, the Nagelian position is characterised by the view that purely cognitive states, such as beliefs, can motivate by themselves without the help of a desire or pro-attitude (1997: 12). What distinguishes Nagel's account, on Brink's view, is the notion that in so far as desires do figure in intentional action, 'they are

merely consequential on our interpreting [an agent's] behaviour as intentional: the action is produced by her moral beliefs and does not depend upon prior, independent conative states' (1997: 13). On the 'consequential ascription' view, the desires in question are more like figures of speech than robust independent mental states. This is not the only way one might interpret Nagel's account, but it seems to many to be the most plausible interpretation.<sup>22</sup>

Dancy notes that proponents of the desire/belief model have often been less than clear about the details of their theory (2000). Specifically, while they all tend to agree that intentional action is the product of two kinds of intentional states with different directions-of-fit – belief and desire – and that desire is always the dominant state, they are unclear about the way in which desire dominates. Here is Brink again, detailing his view of the precise difference between Smith's account, which is meant to adhere to the desire/belief model, and Nagel's account, which rejects the desire/belief model:

On [Smith's] version of rationalism, pro-attitudes are not merely consequentially ascribed because the action is intentional: the pro-attitudes are psychologically real prior to the action and play an ineliminable role in generating, and hence, explaining action. But *these pro-attitudes are consequential or dependent on the belief* that one should perform the action. (Brink 1997: 15. My emphasis)

So on Smith's model of moral motivation it seems that belief rather than desire is the dominant state: the belief brings about or produces the desire. But this suggests that citing a moral belief is sufficient for explaining why an agent is *motivated* to do what she takes to be right. And this is in conflict with Smith's official 'Humean' or desire/belief theory of motivation, according to which motivation is the pursuit of a goal constituted by desire (1994: 104, 116). Consider the following example. John comes to the conclusion that he is morally required to support the local residents' association. On Smith's model, assuming John is not weak-willed and so on, this belief will directly produce a desire to support the local residents' association. If we

---

<sup>22</sup> In *Moral Reasons* (1993), Dancy interprets Nagel's position as something more akin to the model Brink attributes to Smith – that moral beliefs produce robust desires – rather than the consequential ascription model. However, in *Practical Reality* (2000) he adopts the consequential ascription interpretation of Nagel, along the lines of Schueler's (1996) interpretation.

assume that the motivation to support the local residents' association is indicative of the general structure of moral motivation in the virtuous person, then John's motivation must be direct and not merely contingent upon some further desire. The reason John gives for supporting the residents' association is that he is morally required to do so. This moral belief is sufficient, and it is sufficient on Smith's account, to explain why John comes to desire to support the residents' association. The desire is not motivated by some further end or goal at which John is aiming. And this is what appears to be in direct conflict with Smith's claim that motivation is always constituted by a relevant desire and means-end belief (1994: 92, 93).

It looks as if in trying to fill out his specific account of moral motivation, Smith has been forced to abandon the very model of motivation he is concerned to defend.<sup>23</sup> By contrast, the act internalism model I favour and that is espoused by non-cognitivism is perfectly consistent with the desire/belief theory.<sup>24</sup> If part of what an agent is doing when she makes a moral judgement is expressing a motivating attitude, then this attitude can always be cited as the principal source of the motivation.

#### *iv) Reasons and Reliability*

In the last section we saw that the explanatory force Smith attributes to the practicality requirement depends in part on a theory of motivation that is in direct conflict with the theory of motivation he wants to defend. In this section I want to look more closely at the role the notion of a rational requirement or justifying reason plays in Smith's outline of the reliability thesis. I argue that the reliability thesis is

---

<sup>23</sup> This claim requires further support, given Smith's specific contention that his approach is consistent with the desire/belief thesis (1994: 179). I provide this support at the end of (iv), below.

<sup>24</sup> My approach is to establish cognitivist-internalism as a *prima facie* plausible position consistent with the desire/belief thesis. This is intended to address traditional objections to cognitivist-internalism that dismiss it out of hand because it contradicts the orthodox desire/belief theory of action explanation and motivation. However, as will become clear in subsequent chapters, the picture of moral motivation that emerges from the speech act analysis of moral judgements suggests that the desire/belief model is inadequate to deal with the complex structure of moral motivation. In chap 5, I argue that the desire/belief model does not provide an adequate account of the nature of intentional action. In contrast to the approach of other cognitivist-internalists, such as Nagel and Dancy, my critique of the desire/belief thesis is a conclusion of my argument for cognitivist-internalism rather than a premise in my argument for cognitivist-internalism (see Nagel 1970; Dancy 1993, 2000).



not sufficiently theoretically neutral to play the role that Smith wants it to play in his argument.

We need to be clear that the central point at issue between Brink and Smith is *whether there can be failures of moral motivation that do not result from practical irrationality*.<sup>25</sup> This issue is central because they are in agreement about the possibility of genuine moral judgements that fail to motivate. Smith's argument based on the reliability thesis is meant to provide support for the practicality requirement that is *independent* of the assumptions underlying internalism and externalism. Given the point at issue, this means it has to provide an independent reason to support the claim that failures of moral motivation necessarily result from practical irrationality. The problem with Smith's argument is that his statement of the reliability thesis has built into it the suggestion that failures of moral motivation are due to practical irrationalities.<sup>26</sup> It therefore cannot be used as the basis of an *independent* argument to decide the issue between internalism and externalism. Now, the reliability thesis is strictly speaking not question begging with regard to the issue of whether failures of motivation are necessarily practically irrational: it doesn't follow from Smith's statement of the reliability thesis that there is a *necessary* connection between failures of moral motivation and practical unreason. However, because the reliability thesis concerns itself with the 'striking fact' that the connection between moral judgement and motivation is reliable *in the absence of practical irrationalities* it is, surely, far too close to suggesting the very point at issue.

If the reliability thesis is to be used to decide between the practicality requirement and its opponents it needs to be stated in a way that is neutral between the two sides. The fact about moral motivation that I take to be 'striking' is that motivation tracks

---

<sup>25</sup> This is clearly a different question to that with which the debate begins about the connection between moral judgement and motivation.

<sup>26</sup> It becomes clear in Smith's later work that the 'good and strong willed person' is simply an agent who is motivated by her moral judgements in so far as she doesn't suffer from practical irrationalities. He now prefers the term 'moralist' to describe this agent. The moralist is someone who if she 'judges it right to do something then she is motivated accordingly, at least absent practical irrationality' (Smith 1997: 111).

moral judgements in most people most of the time. There is no need to include a *caveat* indicating why or for whom the connection fails to hold. The point is that overwhelmingly it does hold and the question is how to explain this fact. Smith opts to explain the reliable connection in relation to the content of moral judgement. It is important to note that at this stage of his overall thesis, Smith's argument for the practicality requirement is meant to be compelling independently of the specific details of his 'rationalist' theory of moral content. In *The Moral Problem* Smith argues that the practicality requirement does not entail rationalism (1994: 62). Yet it is the assumption that moral requirements are requirements of reason that does all of the explanatory work with regard to the reliability thesis. This comes out more clearly in his later work: 'Remember, I say that the Claim about Moralists is true because of something about the nature of moral judgement: moral judgements are analyzable in terms of beliefs about our normative reasons...' (Smith 1997: 112).

We saw above that the explanatory force of Smith's claim about the content of moral judgements is only as strong as the specific theory of motivation that his model offers. Whereas the externalist explanation of the reliability thesis relies on the widely accepted desire/belief model of motivation, Smith's explanation relies on a controversial theory of moral content. Remember that Smith's version of the reliability thesis asks for an explanation of the fact that in the absence of practical irrationalities agents are reliably motivated by their moral judgements. If moral requirements are requirements of reason as his content claim suggests, then any failures of moral motivation will, plausibly, be practically irrational; so we have a neat explanation of why motivation tracks moral judgements in the absence of practical irrationality. But how does the rationalist theory of content fare in explaining our revised reliability thesis, stripped of all mention of practical irrationalities? Well, if moral judgements are judgements about justifying reasons, and motivation tracks reason judgements, then there will still be a reliable connection between moral judgements and motivation.

The problem for Smith is that once the reliability thesis is revised to exclude explicit mention of practical irrationality, his version of content internalism no longer

provides a simple explanation of why motivation tracks moral judgements. It merely explains one controversial doctrine by way of another controversial doctrine. A much simpler alternative explanation, not involving the externalist model Smith rejects but still plausibly externalist, is that moral requirements normally, but not necessarily, coincide with requirements of reason. On this model, if we accept that motivation tracks reason judgements, then motivation will normally and reliably, though not invariably, track moral judgements. This explanation of the reliability thesis is simpler than Smith's as it doesn't depend on a controversial thesis about the content of moral judgements.<sup>27</sup>

Smith's use of content internalism about justifying reasons to explain the reliability thesis also places him on the horns of a dilemma. This is because in order to explain how judgements about justifying reasons motivate he must either deny the desire/belief model he elsewhere defends or face his own accusation of fetishism. As we have seen, the approach Smith adopts commits him to denying the desire/belief thesis for moral motivation. To avoid this alternative, he would need to retract the claim that reason beliefs, and thus moral beliefs, are capable of producing motivation without relying on a prior goal-constituting desire. However, this leaves him with the task of explaining how, on the desire/belief model, reason judgements or beliefs motivate. According to the desire/belief model, in order to be motivated by a *belief* that I have reason to do some act, there must be some goal constituted by desire that doing what I have reason to do promotes. In order to explain the fact that motivation reliably tracks reason judgements, this goal, it would seem, needs to be something like a standing desire to do what I have a justifying reason to do. But if this were the case, following Smith's reasoning, the virtuous agent would ultimately be motivated by a fetishistic desire to be rational, rather than by a direct concern for the good of others: a misrepresentation of the motivational psychology of virtuous people.

There are a number of features of Smith's theory that might be thought to address the kinds of objections raised above. Consider the following passage:

---

<sup>27</sup> This position together with the act internalist option means that there are at least two alternatives to the practicality requirement that explain the reliability thesis. It is therefore clear that Smith's anti-externalist argument does not immediately vindicate the practicality requirement.

Suppose an agent who does not yet desire to  $\Phi$  deliberates and, as a result, comes to believe that she has a normative reason to  $\Phi$ . And suppose further that her coming to have this belief then causes her to desire to  $\Phi$ . Given C2<sup>28</sup> it follows that we should redescribe this causal transition between belief and desire in normative terms. For her having that belief causes her to have a desire that it is rational to have, given her belief...

Moreover, note the fact that our beliefs and desires may bear such normative relations to each other *is not inconsistent with the Humean [desire/belief] theory of motivating reasons* defended earlier. Indeed, this whole discussion has been premised on the Humean theory. All actions are indeed produced by desires, just as the Humean says; no actions are produced by beliefs alone or by besires.<sup>29</sup> But, if what we have said here is right, some of these desires are themselves produced by the agent's beliefs about the reasons she has, beliefs she acquires through rational deliberation.

We are now in a position to pull the threads of the discussion together. It seems difficult to reconcile the claim that deliberation on the basis of our values is practical in its issue to just the extent that it is with two further claims, the claim that deliberation normally reflects our evaluative beliefs and the claim that our actions are produced by our desires. However, we have seen that these claims are not in conflict. Instead they reflect a substantive fact about human agents: namely, that we are rational creatures who are sometimes more rational, sometimes less. Deliberation on the basis of our evaluative beliefs is practical in its issue to just the extent that it is because *that* is precisely the extent to which we are rational...

For...when we deliberate, we try to decide what we have reason to do, and to the extent to which we are rational we will either already have corresponding desires or our beliefs about what we have reason to do will cause us to have corresponding desires... (Smith 1994: 179-180. My emphasis)

The structure of moral motivation Smith sets up in this passage mirrors that attributed to him by Brink in the passage I quoted in (iii), above. Recall that Brink said that on Smith's model, 'the pro-attitudes are psychologically real prior to the action and play an ineliminable role in generating, and hence, explaining the action. But these pro-attitudes are consequential or dependent on the belief that one should perform the action.' I suggested that this meant that on Smith's model, and contrary to the desire/belief thesis, belief plays the dominant role in motivation. However, Dancy has suggested that we should distinguish between causal dominance and Humean dominance (2000: 84). While the belief in Smith's scenario may be

---

<sup>28</sup> Smith (1994: 148): 'C2 If an agent believes that she has a normative reason to  $\Phi$ , then she rationally should desire to  $\Phi$ '.

<sup>29</sup> 'Besires' are psychological states said to have both belief-like and desire-like directions of fit. In this way it is thought they might capture the cognitive and the motivating aspects of moral and normative judgements (see Lewis 1988; Price 1989).

*causally* dominant, the desire can still play the dominant role in motivating action, a role it has due to its direction of fit.<sup>30</sup>

Smith stresses that his model of moral motivation is consistent with the desire/belief thesis because it is consistent with the view that all action is produced by a combination of belief and desire and not simply by belief alone. But my objection is based on the fact that Smith presents the desire/belief model, or as he terms it the ‘Humean theory of motivation’, as a theory of *motivation*, and not simply a theory of action explanation (1994: chap. 4). And he tells us that, on the Humean theory, *motivation* is the pursuit of a goal constituted by desire (Smith 1994: 104, 116).

To see that the theory of motivation and the theory of action explanation are distinct, consider again the example of John’s being motivated by the consideration that he is morally required to support the local residents’ association. Let’s say that John has not yet acted on his decision to help. This allows us to focus on why he is motivated to help the local residents’ association and not his action of helping them. What we need to explain is why John comes to be motivated to help the local residents’ association. According to Smith, the Humean theory of motivation tells us that John’s motivation must be explained in terms of his pursuit of a goal constituted by desire. But this is not how Smith explains moral motivation; rather, he says that desires like John’s are *caused* by moral beliefs. This allows Smith to avoid the fetishism horn of the dilemma: John is not motivated to do what he judges is right in

---

<sup>30</sup> Dancy suggests that this point undermines Nagel’s attack on the desire/belief model based on the latter’s distinction between motivated and unmotivated desires (2000: 84). This might be true if two things were the case: if, as Dancy suggests, Nagel’s conception of a motivated desire is of a desire that is caused or produced by a belief state; and if the desire/belief model only purported to provide necessary conditions for the explanation of intentional action, and did not also purport to be a theory of motivation. I deny both of these claims. As I explain in chap 5, the concept of a motivated desire is the concept of a desire held for a reason or on the basis of deliberation, not the concept of a desire merely caused or produced by a belief state. This means that Nagel’s conception of a motivated desire implies a conception of motivation radically different to that assumed by most proponents of the desire/belief thesis. In chap 5, I call this kind of motivation, normative motivation. As I explain in the text above, I also reject the claim that the desire/belief thesis is standardly proposed as an account of action explanation and not, in addition, an account of motivation more generally.



order to satisfy some goal of his or to achieve something he desires to do.<sup>31</sup> But in avoiding the fetishism charge, Smith's approach is caught on the other horn of the dilemma: it contradicts that claim that motivation must always be explained in terms of a goal constituted by desire.

Smith might deny that the desire/belief theory is meant to be a theory of motivation as well as a theory of action-explanation. But if the desire/belief model is restricted in this way it becomes a much less interesting and much less contentious thesis. For example, it would make it broadly consistent with the anti-Humean theory Dancy calls 'pure cognitivism'.<sup>32</sup> Perhaps Smith's theory differs from anti-Humean theories like pure cognitivism in ultimately explaining motivation and reliability in terms of background conative dispositions (Smith 1994: 180)<sup>33</sup>. Smith refers to this aspect of his theory in the passage quoted above. His point seems to be that rational agents are agents who, at a minimum, are disposed to do what they judge they have good reason to do. It is part of being a rational agent that if one believes there is an overriding reason to do something, then, *ceteris paribus*, one will desire to do it. This is an interesting and promising suggestion, but it takes us a long way from the notion that motivation is the pursuit of a goal constituted by desire.

---

<sup>31</sup> I argue in chap 5 that it is not clear on this type of model that the agent is motivated at all in a sense consistent with the concept of motivation involved in intentional action. On Smith's model, John's desire to help does not seem to be motivated by his taking its rightness as a consideration in favour of helping; rather it seems to be causally, and non-rationally, produced. The desire is, admittedly, *consistent* with what it is rational to do, but it is 'merely the result of the operation of [the belief] *in* [him] rather than an expression of [his] own mental activity' (see Korsgaard 1997: 221; and 5.4, above)

<sup>32</sup> See Dancy (2000: 90). Pure cognitivism accepts that in order for action to take place there must be both a belief and a desire, and that these are distinct states with different directions of fit. However, it sees belief as the state which motivates and desire as the state of being motivated. In chap 5.3, I argue that beliefs are states that convey considerations that motivate, whereas desires are, normally, states of being motivated.

<sup>33</sup> For Smith this appears to involve our being disposed to having a 'systematically justifiable set of desires', where the process of systematic justification is akin to Rawls' concept of reflective equilibrium (Smith 1994:159; Rawls 1973: 48-51). Actually, Smith describes this disposition as our having 'a goal of having a systematically justifiable set of desires' (1994: 160). However, it cannot be the goal of having a coherent set of desires that motivates agents: this risks a form of desire-coherence fetishism. I take it that we would be unimpressed by an agent who declined to help because he was worried that to do so would make his desire-set less coherent. If the disposition is a background condition that makes agents susceptible to normative considerations then, again, it suggests a far more complex theory of motivation than that described by the desire/belief model.



The notion that agents have general conative dispositions which make them susceptible to motivation by normative considerations is an idea I develop in chapter five. But the structure of moral and normative motivation it suggests is far more complex than can be accommodated on the standard desire/belief model. In the next section, I argue that if there is such a thing as genuine moral motivation, then it must be possible for agents to be motivated by the *consideration* that something is morally right.<sup>34</sup> The notion of being motivated by favouring considerations is something absent from most versions of the desire/belief thesis. If moral motivation is sometimes non-instrumental, and thus non-fetishistic, it must be possible for agents sometimes to do what is right because it is right, rather than because doing so satisfies some further goal or desire. But this is something that cannot be accommodated by the desire/belief thesis if it is intended as a theory of motivation as well as a theory of action explanation.

v) *The Accusation of Fetishism*

Smith thinks that those who reject explaining the reliability thesis using the practicality requirement must explain it in terms of some form of general or standing desire to do the right thing. He claims that externalism misrepresents the motivational psychology of virtuous people because virtuous people care directly for what they judge as right, not derivatively via an independent desire to be moral (Smith 1994: 75). I have already argued that there are at least two other options available for explaining the fact that motivation reliably tracks moral judgements: It might be a necessary condition of performing a genuine moral judgement *act* that an agent who makes a judgement is motivated to conform to it. Alternatively, if moral requirements are generally, though not invariably, judged to provide good reasons for action, then, given that motivation tracks reason judgements, moral judgements will reliably motivate those who make them. I leave these considerations for the moment in order to focus on Smith's objection to the standing desire model of externalism.

The argument I presented in section two agrees with Smith that externalism must adopt a standing desire model of moral motivation in order to accommodate the way

---

<sup>34</sup> See chapter 5.3 for an explanation of the notion of a favouring consideration.

in which motivation tracks moral judgements. Smith rejects the standing desire model because, he argues, it makes the motivations of virtuous people fetishistic; and this is something that cannot be true of virtuous agents. He views the fact that externalism is committed to this false view as a *reductio* of the externalist position. According to Smith, externalist agents are ultimately motivated to do what is right ‘for the sake of its being the right thing to do, rather than for the sake of the very feature that they believe makes it the right thing to do’ (Smith 1997: 114). He presents this distinction in terms of *de dicto* and *de re* desires to do what is right. If I have a *de dicto* desire to do what is right, then I desire to do what is right *because* it is the right thing to do; whereas, if I have a *de re* desire to do what is right, then I desire to do what is right *because*, for example, someone is in need or because doing so will prevent harm: the features that make it the right thing to do (Smith 1994: 74). Smith says that the externalist model cannot accommodate the fact that virtuous people care directly for the well-being of family and friends, honesty, and other goods, and not instrumentally in so far as their acts satisfy the goal of doing what is right. The consequence of this is that the externalist model ‘alienates [agents] from the ends at which morality properly aims’ (Smith 1994: 76).

There are a number of different objections Smith makes in advancing his fetishism argument that need to be identified to make sense of his concerns. One consideration Smith finds troubling is that the externalist model generates ‘derived’ desires to do those things that are deemed right, and that such desires are ‘underived’ in the virtuous person. Presumably, Smith is not claiming that the virtuous agent’s desires are all unmotivated, in the sense that they are not based on any other considerations or not held for any reasons. This would make such desires underived, so that, like hunger or thirst, they simply assail the agent rather than being arrived at via rational deliberation. There is a tradition that views virtuous agents as those who are habituated to do what is right so that much of their behaviour is of the appropriate kind unmediated by overt deliberation. For example, a kind person will respond to someone in need when she becomes aware of their need without having to engage in complex moral deliberation prior to acting. Nonetheless, if her action is rational she will be motivated to help for a reason: she won’t simply be assailed by the desire to

help. Given Smith's rationalism, I think we can assume that he doesn't hold that moral desires are underived in the sense of not being held for a reason.

Sometimes Smith expresses his worry about externalism in terms of 'instrumental' and 'non-instrumental' desires (1997: 114-115). Presumably, again, the distinction is not meant to correspond to the that between desires held for a reason and those held for no reason. If an agent desires something instrumentally, this normally means that she desires it for the sake of something else: her reason for desiring it is in order to satisfy some further end or purpose. By contrast, if an agent desires something non-instrumentally, this normally means that she desires it for its own sake. Smith argues that externalism is committed to the view that agents only desire instrumentally to do those things that are in fact right – looking after family and friends, and other goods – in so far as doing such things satisfies the goal of doing what is right. The virtuous agent, by contrast, desires to look after family and friends and do other good things non-instrumentally: she has non-instrumental concern for family and friends and those in need.

An interesting feature of Smith's objection based on the distinction between instrumental and non-instrumental motivation is that it potentially generates a problem for *any* theory of moral motivation. His fundamental objection is not that the source of the agent's motivation is a *desire*, but that the content of the desire, taken as the reason for acting, is 'to do the right thing.' But a purely internalist and cognitivist view of moral motivation potentially faces exactly the same charge.<sup>35</sup> For example, suppose I judge that I am morally obliged to volunteer to help at the local charity shop. Let's say that Smith's model of moral motivation is correct, so that my belief that it is right to volunteer produces 'straightaway' a desire to do so. The first thing to notice is that my desire to help *derives* from my belief that it is right to volunteer. Now, presumably my desire is held for a reason and does not merely assail me. What reason might I give to explain my motivation? Plausibly, I will

---

<sup>35</sup> A suggestion by Richard Holton lead me to think about this style of response to Smith's fetishism charge.

desire to volunteer because I believe it to be the right thing to do. But this now looks like my desire to help is merely a means to achieving the end of doing what is morally right. So it appears that Smith's cognitivist model of motivation is also vulnerable to the charge of fetishism.

Another example might help to understand what has gone wrong with Smith's argument. Suppose that one were genuinely motivated to work for someone because they were the President. In such a case, the consideration that they were President would have to be the consideration or reason in light of which one desired to work for them. We might call such a desire a *de dicto* desire to work for the President. A *de re* desire to work for the President might take the following form. If Carter is the President, then John desires to work for the President because John desires to work for Carter. However, to be genuinely motivated by the fact that someone is President, it must be the person's status as President that is the relevant motivating consideration. If we were to analyse genuine instances of 'motivation-by-presidency', we should limit our analyses to those cases in which an agent is motivated by the consideration that someone is the President. Although it is true that John is motivated to work for the President, it is not the case that he is motivated by the consideration that Carter is President; rather, he is motivated by the consideration that the President is Carter. So, given our criterion, John's being motivated to work for the President is not a genuine case of motivation-by-presidency.

If it were not possible to be motivated by the consideration that someone was President rather than that they were a particular person who happened to fill this role, then there would be no such thing as motivation-by-presidency. It seems to me that we have an analogous situation in the case of moral motivation or 'motivation-by-rightness'. If there is such a thing as moral motivation then it must be possible for agents to be motivated by the consideration that something is morally right. If John is motivated to help the poor in order to alleviate their suffering, we might say that John is doing something that is in fact morally right. But unless John is helping the poor and alleviating their suffering *because* it is morally right, then we do not have a genuine case of moral motivation; we just have an act that happens to accord with

morality. Or suppose that John is motivated to do what he judges is right *because* it involves volunteering. This would fail the tracking condition. If it became clear that the right option didn't involve volunteering, we have no reason to think that John would lose his motivation to volunteer and take the moral option.

We can make the same point from a different perspective if we take as an example an act that is not obviously admirable. Headmaster Watt is motivated to give boys who get bad grades the cane. If Watt judges that he has a moral obligation to beat boys who get low grades, and this is what motivates him to cane them, then we have a genuine case of moral motivation, regardless of how morally mistaken we take Watt to be. But if Watt doesn't construe his reason in moral terms (perhaps he just wants to improve the school average), then we have no reason to interpret his motivation as moral motivation. Smith, in arguing that motivation by a consideration that something is right *de dicto* is necessarily fetishistic, threatens to make moral motivation impossible. And surely this is a *reductio* of his argument.

Another strand of Smith's fetishism objection is that the externalist agent is unhealthily obsessed with the moral status of all her actions. I think we can agree that a virtuous person is someone who knows when it is and is not appropriate to engage in moral deliberation and justification. But there is nothing about the structure of the externalist position that compels the externalist agent to engage in moral deliberation about each and every decision she makes. She need not be fetishistic in the sense of being constantly obsessed with the moral status of her actions. This strand of Smith's objection perhaps explains his appropriation of an argument by Williams designed to show the limits of an impartial morality. But Smith misapplies the Williams objection, which is about when moral judgements are and are not appropriate sources of motivation, not about the structure of motivation when moral judgements do motivate. Williams discusses a situation in which a man is faced with saving the life of one of two drowning people in equal peril, one of whom is his wife and the other a stranger (1976: 17-19). One question is whether the special relationship the man has to his wife provides a moral justification for his treating her differently to the stranger: when is it fair to give preference to one's

family and friends over strangers? Williams, however, is concerned about the appropriateness of engaging in moral deliberation *at all* in such situations: thinking about the moral status of his actions in such a context rather than simply the fact that the person drowning is his wife ‘provides the agent with one thought too many...’ (Williams 1976:18).

For the sake of argument, let’s agree with Williams that in the case he describes it is inappropriate for the husband to appeal to moral justification at all. So when the man is motivated appropriately to save his wife he is motivated by the thought that it is his wife who requires saving. For Williams, then, this kind of motivation lies beyond moral justification and is not an example of moral motivation. Now, suppose that the husband were morally motivated according to Smith’s model. In this case, his belief that it was the right thing to do would produce a desire to help his wife. But it is precisely this kind of dependence on moral considerations that Williams says involves ‘one thought too many’. And this charge potentially applies to any model of moral motivation. There is no reason to think that an externalist, any less than an internalist, would be unable to discern those cases in which moral considerations do not apply and in which, therefore, moral motivation is not an issue. Again, it seems that a concern Smith raises for externalism is equally applicable to internalist theories of moral motivation.

## **5. Externalism and the Two Views of Morality**

I concluded in the last section that Smith’s fetishism objection is ultimately unsuccessful in undermining externalism. However, there are other reasons for being sceptical about the externalist thesis. Much of the impetus for externalism is undermined once we dismiss the non-cognitivist anti-cognitivist argument. If cognitivism and internalism are not incompatible in the way that externalism fears, then, given a number of significant problems with externalism, I think we have reason to prefer some form of act internalism. In this section, I argue that externalist explanations of the practical or action-guiding function of morality are unconvincing. If motivating force and practical normativity are only contingently related to moral considerations, then it is difficult to account for the role such considerations play in



our practical lives. I also claim that if we focus on the function or point of moral practice, we can see why someone outside the practice, such as the amoralist, will be unable to make genuine moral judgements.

Brink is clear that his preference for moral judgement externalism is largely a result of the threat he believes internalism poses to realism in ethics. Moral realism and internalism are, he says, 'uneasy bedfellows' (Brink 1989: 43). Much of this threat he thinks derives from the non-cognitivist anti-cognitivist argument (Brink 1989: 44). Shafer-Landau refers directly to a version of the anti-cognitivist argument to explain the view that moral judgement internalism and moral cognitivism are incompatible: 'necessarily, moral judgements motivate. Beliefs don't. Therefore moral judgements aren't beliefs' (2000: 270). But if, as I have claimed, the anti-cognitivist argument is invalid and cognitivism and internalism are *prima facie* compatible, then much of the initial impetus for externalism is removed. This doesn't show that externalism gives the wrong account of moral judgements, but, when added to other difficulties with the externalist thesis and the intuitive pull of internalism, it does significantly weaken the case for externalism.

Externalists generally accept that moral considerations have an important role in guiding action and in practical deliberation more generally: 'Moral considerations are practical in some very important sense. Agents engage in moral deliberation in order to decide what to do and give moral advice...' (Brink 1989: 37). Shafer-Landau makes a similar point: 'Fact: everyone we know is motivated to some extent to comply with his or her moral judgements. Fact: we suspect the sincerity of someone who proclaims fidelity to a moral code, all the while showing no inclination to abide by it' (2000: 284). Shafer-Landau suggests that despite appearances such considerations are neutral between internalism and externalism. This is because while agents are usually motivated by their moral judgments, we don't observe every person as motivated to comply with each of her moral judgements. So it seems like everyone is always motivated to some extent to comply with their moral judgements but only because in most cases they are motivated to do so.

Svavarsdóttir also notes the force of the connection between moral judgement and practicality:

Admittedly, I would probably experience puzzlement upon encountering an individual who made a moral assessment of his circumstances yet appeared indifferent to moral values...Why would anyone who is completely indifferent to moral considerations bother to take note of the moral conditions of his surroundings? (Svavarsdóttir 1999: 184-185)

Svavarsdóttir argues that this puzzlement can be explained by the general assumption we make that someone who bothers to make a moral judgement is to some extent committed to morality. If commitment implies motivation then, given the assumption about moral judgement and commitment, we will expect agents to be motivated to comply with their moral judgements.

By their own admissions, it is incumbent upon externalists to account for the 'important sense' in which morality is practical. If we look at the intuitions they mention above there is a common theme: the point or function of morality appears to be tied to practical deliberation. Shafer-Landau tries to explain this away by appealing to the consideration that agents are usually but not always seen to be motivated by their moral judgements. Svavarsdóttir suggests that moral judgements seem necessarily to be tied to decision making only because we assume people who make them are morally committed.

It is very difficult for externalism to accommodate the apparently close connection between moral judgement and practical deliberation. This is because moral judgement externalism holds that the motivational force of moral judgments is entirely dependent on contingent desires to do actions of the relevant kind. And according to content externalism about justifying reasons, moral judgements and considerations only ever contingently justify. If one accepts both these forms of externalism, then one must accept that moral judgements and moral considerations have no inherent normative authority and no inherent motivating power. This seems at odds with the function of such considerations and judgements being to motivate action and to guide choice. If moral judgements do not have an essentially action-guiding function, then it seems somewhat mysterious that moral considerations

overwhelmingly and consistently play the role they do in our practical lives. What we refer to as ‘morality’ and as ‘moral codes’ in different societies and in different historical periods appear to have as a central function the regulation of conduct. The externalist needs to provide a convincing explanation of this apparent fact.

Is the inherent practical impotence of moral judgements and considerations on the externalist model a problem for externalism? To answer this question we need to reflect on the ways in which externalism attempts to account for the practicality of morality. We saw in the last section that one way in which Brink attempts to account for the motivating force of moral judgments is via the notion of sympathy. I argued that non moral-content versions of externalism of this kind fail to account for the way in which motivation tracks moral judgements. Agents are motivated on these models by desires to do the kinds of things morality is purported to be about. For example, if morality is about contributing to the general good, moral motivation is explained by desires agents have to help others, or perhaps by other desires that are satisfied by contributing to the general good. This approach is conducive to the common externalist aim of defending realism and naturalism in ethics: moral facts are less likely to be considered metaphysically queer or epistemologically inaccessible if they are simply facts about what contributes to the general good or to human happiness. Moreover, it seems obvious that someone could become aware of such considerations and remain indifferent to them. For example, I might recognise that morality requires me to help at the school fundraising event, but desire instead to stay at home and read.

Considerations of the above sort bring us very close to the central issue that divides externalists and internalists. Internalists, typically, do not characterise moral judgements in terms of their content but in terms of their function or role. Brink, in a footnote, spells out this crucial point:

Sometimes the debate between internalists and externalists depends on the two employing different criteria for identifying judgements as moral judgements. Often, internalists employ functional criteria – roughly, those judgements are moral judgements that the appraiser treats as fundamentally important – whereas externalists employ contentful criteria – roughly,

those judgements are moral judgements that concern certain sorts of matters, for instance, having to do with the welfare of affected parties. (Brink 1997: 21)

Brink says little more on this point other than that he is committed to rejecting the functional criterion (1989: 22). Mackie refers to what I take to be essentially the same distinction (Mackie 1977: 106). A ‘broad conception of morality’, he argues, is any general theory of conduct that helps to direct an agent’s choices. Morality in its ‘narrower sense’ suggests a specific set of other-regarding concerns centring around the well-being or interests of human agents. Mackie makes the point that while we cannot help but act on, or have a perceived reason to act on, what we take to be required on the broad view, it is quite possible that our preferences might run counter to the dictates of morality in its narrow sense: ‘I admit that morality requires that I should do such-and-such, but I don’t intend to: for me other considerations here overrule the moral ones’ (Mackie 1977: 106).

The distinction between the broad and narrow views of morality obviously has direct bearing on the debate between internalism and externalism. On the broad or functional view, moral judgements are intimately tied to practical decision making in a way that suggests an internal or conceptual connection to motivation. By contrast, on the narrow or content-based view, whether I am motivated to act morally will be contingent upon my having a commitment to the kinds of actions and concerns that characterise the narrow view. This suggests that there is a much deeper question about the nature of morality that underpins the debate between externalists and internalists. However, Mackie at least thinks that it is difficult to privilege one view over the other: ‘Both are used and both have important roots and connections in our thought’ (1977: 107). If Mackie is correct, then we may have an important explanation of the apparent ‘stalemate’ between internalists and externalists: the intuitions of both sides are attuned to equally important *but different* facts about ordinary moral thought.

While I accept that the intuitions of both sides have their source in significant aspects of ordinary moral thought, I suspect that the narrow view reflects widespread consensus about answers to questions raised by the broad view – about what is

ultimately practically valuable or important – and that, as such, it is less fundamental than the broad view. The narrow view, on this interpretation, reflects one kind of first-order answer to questions of value. The amoralist is always a real possibility on the narrow view, because it is always possible to lack interest in or question the justification of any particular considerations construed as moral. By contrast, if I judge that something is morally right on the broad view, I make a judgement about what is ultimately practically justified or important. It makes no sense on this view to question the practical justification of one's moral judgements as they are simply judgements about what is ultimately practically justified.

The issue of whether we should view morality primarily in terms of the broad or the narrow view deserves a level of consideration that is beyond the scope of this work. In developing my positive account of moral judgement I adopt the broad reading of morality, as described by Brink. On this account, morality has to do with what is of fundamental importance in terms of agents' conduct, and moral judgements are those judgements that concern what an agent takes to be of fundamental importance in this respect. My approach is to adopt the broad view of morality to determine if it can provide the basis for a plausible version of cognitivist internalism. I will not defend the broad view over the narrow view directly. Nonetheless, I hope to show that the model of moral judgement it supports makes such a view attractive.

Hare is wrong that amoralists merely describe what other people think is right. The amoralist externalists have in mind believes herself to be making genuine moral judgements, and on the narrow view, a view with significant historical and philosophical credentials, she does make such judgements. However, what interests Hare is morality in the broad sense: the general principles of conduct that function to guide action (1952: 1). And in this sense, the amoralist sceptic, at least, doesn't seem to be making genuine moral judgements because she doesn't accept the practical normative authority of her judgements of moral rightness. The unprincipled amoralist, who doesn't question the normative authority of her moral judgements but remains unmoved by them, is a more complex matter. The internalist needs to explain how moral judgements can both fail to motivate and have an essential

connection to motivating attitudes. And act internalism must explain this without including the relevant conative or motivating states at the level of judgement content. The speech act approach I develop in chapters three and four will provide the framework for a version of act internalism that can account for failures of motivation in this way.

Svavarsdóttir, in her defence of moral judgement externalism, states that internalism would be an obvious conceptual truth if moral commitment were a precondition for making genuine moral judgements (1999: 185). She argues in favour of externalism that agents can fail to have such a commitment and still make moral judgements: The ‘commitment sceptic...does not dispute that there are morally better and worse alternatives, but wonders why that should affect our decision making and action’ (Svavarsdóttir 1999: 183). On the broad view, the commitment sceptic is hopelessly confused because the point of moral deliberation is to find out precisely those things that should affect our decision making and action. In chapter one, I said that internalism reflects the participatory nature of moral judgement. So we might say that there is a sense in which commitment to morality is a precondition for making genuine moral judgements. That is, one must be a participant rather than an onlooker. And this is why the amoralist fails to make genuine moral judgments: she is, by her very nature, outside morality, at least in its broad sense.

The first two chapters have indicated a far more complex picture of moral psychology than we find in most standard internalist and externalist accounts. Specifically, motivation and motivating states seem to figure in moral deliberation *at a number of different levels*. I think that this reflects the ambiguous nature of the term ‘motivation’. I suggested in chapter one that ‘motivation’ can refer to the psychological state of being motivated, the consideration in light of which an agent is motivated, or to the psychological state which conveys the consideration in light of which an agent is motivated: respectively, a desire, an agent’s reason, and a justifying belief with a reason claim as its content. Externalist explanations of motivation tend to focus on the role of motivating *states*. Sympathy, for example, involves a state that disposes people to act to help others. But the fact that an agent



is sympathetic is rarely, if ever, the consideration that motivates the agent to help someone else. This would more likely be the fact that the person in question is in need of assistance.

Although the standing desire model posited by some externalists is able to account for the way in which motivation tracks moral judgements, it suggests, like sympathy, an implausible or inappropriate ground for doing what is morally required. The problem with this kind of externalism is not that its agents are motivated by a *de dicto* desire with the content, 'to do what is right', but that they are motivated by the *desire* to do what is right rather than the consideration *that it is right*. Reformulated in this way, I think that Smith's fetishism objection succeeds in casting doubt on externalism. We expect virtuous agents to be motivated by the consideration that something is right and not by the fact they happen to desire to do what is right. On the latter model, the fact or consideration that something is morally required would only ever be instrumental in satisfying the desire to do what is morally required. This difficulty for externalism may be part of a more general problem for adherents of the desire/belief model (cf. Dancy 2000; Schueler 1995).<sup>36</sup> If so, Smith's adherence to the desire/belief model explains why he is not in a position to formulate the objection in this way.

## 6. Conclusion

Externalists argue that the possibility of amoralism undermines the internalist claim that there is a connection of conceptual necessity between moral judgements and motivation. However, internalists and externalists have opposing intuitions about purported examples of amoralism, making standard thought experiments inconclusive. Smith offers an objection to externalism that he argues is independent of the assumptions underlying the two sides in the debate. However, Smith's statement of the reliability thesis suggests the very point at issue between his version

---

<sup>36</sup> The source of the difficulty being that on the desire/belief model an agent can never be motivated by the consideration that favours action but only, ultimately, by the desire to perform the action. I suspect that this difficulty stems from proponents of the desire/belief model's not distinguishing clearly between motivating states and motivating considerations. At times the desire/belief theorist seems committed to denying the very possibility of motivation by normative reasons (cf. Korsgaard: 1986).

of internalism and externalism. It cannot therefore function as the basis of an independent argument against externalism. Moreover, Smith's version of internalism depends for its explanatory force on a form of content internalism about justifying reasons rather than a version of judgement internalism about motivation. The connection Smith posits between this kind of content internalism and motivation depends on a theory of moral motivation that is in direct conflict with his adherence to the desire/belief model. Act internalism, by contrast, can accommodate the conceptual connection between moral judgements and motivation along with cognitivism in a way consistent with the desire/belief model. But we may wonder whether we should accept the desire/belief model given that, ultimately, it seems to be a weakness rather than a strength of externalism. Although Smith's fetishism objection as stated is unsuccessful, it can be reformulated in a way that undermines the plausibility of externalism. On the externalist model, an agent can only ever be motivated to do what is right in so far as doing so satisfies an independent and contingent desire to do the right thing. And this does seem to conflict with our notion of what motivates the virtuous person. In the next chapter, I apply speech act theory to the intuitions that support act internalism in order to develop an approach that can accommodate the complex moral psychology that has been suggested by the enquiry thus far.

---

## Speech Acts and Moral Motivation

### 1. Introduction

I have argued so far in favour of interpreting moral judgement internalism as a thesis about judgement acts rather than judgement contents. Broadly speaking, act internalism is the kind of internalism espoused by non-cognitivism. Non-cognitivism emphasises the way in which agents are typically *doing* something more than issuing a bare description when they make moral judgements, such as endorsing or commending what is judged right. Thus far I have assumed that the distinction between judgement act and judgement content is clear enough for us to see the implications of opting for act internalism rather than content internalism, or *vice versa*. In order to avoid complicating these discussions unnecessarily, I have not as yet given a specific account of the nature of so called judgement *acts*. To an extent this lack of specificity reflects the way in which the notion of moral judgement is used in the literature on moral motivation.<sup>1</sup> In addition to the structural ambiguity between act and content already discussed, the notion of moral judgement is often ambiguous between moral mental states and moral utterances or speech acts.

My aim in this chapter is to explore the possibility of elucidating issues to do with internalism and moral motivation by approaching the problem from the perspective of moral language. Of course, focussing on moral language may merely serve to emphasise the primacy of moral psychology in these debates. However, I believe that applying speech act theory to the issues discussed thus far will elucidate the structure of moral judgements and their relation to moral motivation more clearly than an account that begins with an emphasis on moral psychology. And I think we should take seriously an approach to these issues that can accommodate many of the purportedly conflicting intuitions underlying our notion of moral judgement. The

---

<sup>1</sup> Some theorists clearly think of moral judgements primarily in terms of linguistic acts (e.g., Stevenson 1944; Hare 1952). More recently, theorists have emphasised moral psychology rather than moral language in discussions of moral judgements (e.g., Nagel 1970; McDowell 1985; Dancy 1993; Smith 1994).

contention of the speech act approach is that an understanding of how and why we are moved by moral considerations can be reached, at least in part, by understanding how and why we use moral language.

In section two, below, I introduce the speech act approach, and in section three, I show how Searle's speech act theory provides a structured way of distinguishing between the act and content of a moral judgement in terms of the difference between a judgement's illocutionary force and its propositional content. Next, I enquire whether the internalist requirement can be met via the conditions that govern the linguistic act of assertion. I conclude that none of the conditions governing assertions entails that the speaker holds a particular motivating attitude or desire towards the propositional content of an assertive utterance. This means it is unlikely that the internalist requirement can be met by interpreting moral judgements as bare assertions.

I then suggest that moral judgement internalism can be accommodated using the notion of a compound speech act. A compound speech act allows speakers to perform more than one linguistic act using a single utterance. I argue that we can meet the internalist requirement on moral judgements by establishing that their performance in speech situations invariably involves the performance of linguistic acts linked to motivating states. In order to understand the different acts involved in making a moral judgement, I suggest thinking of such judgements in relation to the compound act of moral *evaluation*. After a brief discussion of Searle's theory of illocutionary acts, I then outline the conditions I take to be necessary for the successful and sincere performance of a moral evaluation. I argue that both the fact-stating and the motivating aspects of moral judgement can be met via these conditions.

In section seven, I discuss an approach to the issues of internalism and moral motivation by David Copp (2001). Copp's approach is similar to mine in that he attempts to accommodate moral judgement internalism and moral cognitivism by construing the internalist requirement in terms of the linguistic conventions

governing the performance of moral judgements. I argue that there are a number of reasons for preferring my account to his. I then address the problem of adequately distinguishing between the speech act involved in making a moral judgement and the meaning of moral predicates used in making such a judgement. I argue that it is important not to confuse how evaluative predicates are used with their literal meaning, which must remain constant across different grammatical contexts. The particularly close relation between evaluative predicates and the act of evaluation is then explained in terms of the role of evaluative predicates in the institution or practice of evaluation. I end the chapter with a brief review of the arguments of the first three chapters and an outline of some advantages and potential disadvantages of the speech act approach to moral judgement.

## **2. Moral Motivation and Moral Language**

In chapter one, I noted that non-cognitivists focus predominately on the act of calling or saying that something is morally right and the point or function of moral language (e.g., Hare 1952: 169; Gibbard 1990: 6). This focus means that their particular account of moral psychology derives from a theory about what agents are doing when they issue moral utterances. Non-cognitivists argue that if issuing a moral judgement entails motivation, then, because no descriptive or fact-stating claim entails motivation, moral judgements must involve something other than a descriptive or fact-stating claim. I rejected the conclusion of this style of anti-cognitivist argument and offered the following alternative:

C(ii) Judging that an act is morally right either involves something other than a judgement that some fact or state of affairs obtains or something in addition to such a judgement.

This is an important result because our ordinary experience of morality suggests that moral judgements involve a fact-stating component and an affective or motivating component. So a theory that can accommodate both these features is preferable to one which forces us to revise our ordinary thinking about morality and accept only one of the features (cf. Smith 1994: 13).

In order to defend the second disjunct of the revised conclusion, we need to understand the nature of the affective addition and its relation to the fact-stating component of moral judgement. But many standard approaches to language appear to block the possibility of combining fact-stating and motivating features in a single judgement or utterance. For example, according to non-cognitivism the gulf between factual and practical concerns is enshrined in our very use of language. So while some parts of our language are fact-stating, other parts are used to *do* things such as issue commands or express feelings (see Stevenson 1937; Hare 1952: esp. chap. 5). The task then becomes one of determining whether a particular piece of language is fact-stating *or* performative. However, because our use of language is not always consonant with the grammatical form in which we express ourselves, linguistic function can be difficult to determine. For example, the standard linguistic device used to make statements is the indicative sentence, but not all utterances we make using indicative sentences should be thought of as assertions of fact. Specifically, according to non-cognitivism, evaluative utterances are best understood as imperatives or expressions of feeling and not as the truth-evaluable statements suggested by their surface grammar.

In *How to Do Things with Words*, John Austin notes this common distinction between fact-stating language and language that seems to involve our *doing* something rather than just saying something (1962). Statements, assertions, descriptions and the like, according to this view, are distinguished by their assessability in terms of truth and falsity and their association with belief and objective fact. In contrast, performatives such as commands, exclamations, promises, requests, namings and the like are not assessable in terms of truth and falsity but in terms of how well they are performed. If I say to you, 'Shut the window', it makes little sense to ask whether what I have said is true or false, but it does make sense to ask if the utterance was appropriately performed. For example, my choice of expression would clearly be inappropriate if it were obvious that the window was already closed. Austin's profound insight was to note that the utterance of statements, descriptions, assertions and the like, just as much as the utterance of so called 'performatives', involve the speaker's doing and not just saying something.



Speaking a language, he argues, has to do with performing linguistic or speech acts according to certain rules. And this is the case regardless of the type of sentence used in the performance of the speech act. Statements, as well as imperatives and exclamations, are assessable in terms of how well they are performed. Austin's insight points to a way of understanding how we combine practical and factual elements in our use of language. If Austin is correct, then the combination of the factual and the practical, seemingly typical of our use of moral expressions, is ubiquitous rather than anomalous in our linguistic practice.

In order to explain how speech act theory can help our understanding of moral judgements it needs to be examined in more detail. To do this I make use of Searle's theory of the nature and structure of speech acts and make use of his terminology.<sup>2</sup> Searle's theory provides the basis for a systematic account of the relations between moral utterances and the various functions such utterances are used to perform. I argue that this allows us to accommodate a range of intuitions central to the notion of moral judgement.

### **3. The Structure of Speech Acts: Illocutionary Force and Propositional Content**

Like Austin, Searle holds that speaking a language involves the performance of acts according to certain rules. And understanding how agents communicate involves understanding what a speaker is doing when she 'seriously' (not play acting) and 'literally' (not ironically) utters a sentence to a hearer or hearers (1969: 57).

According to Searle, in most cases the serious and literal utterance of a sentence will involve the performance of the following acts: the act of uttering certain words or '*utterance act*'; the acts of referring and predicating, together called the '*propositional act*'; and an act such as stating, questioning, promising, wishing, requesting, and so on, or '*illocutionary act*' (1969: 24). These different acts are meant to be viewed as abstractions from the total speech act. While an utterance act can be performed independently of a propositional act or an illocutionary act – for

---

<sup>2</sup> Searle extends and adapts the work on speech act theory begun by Austin. See Searle 1969; 1979; 1989.

example, I can say words without expressing a proposition or stating, questioning, promising, etc. – a propositional act can only be performed if an utterance act and an illocutionary act are also performed. For example, I cannot refer to a particular window and predicate the expression ‘is open’ of it unless I *state* that the window is open and *utter* the sentence, ‘the window is open’. Similarly, the performance of an illocutionary act necessarily involves the performance both of an utterance act and a propositional act. I cannot state or question or demand that the window is open unless I utter the appropriate sentence and perform the appropriate propositional acts of referring and predicating.

For our purposes the distinction between illocutionary acts and propositional acts is worthy of special attention. Searle takes as the grammatical correlate of propositional acts sentence-parts rather than complete sentences (1962: 29). For example, the statement ‘Sam smokes habitually’, the question ‘Does Sam smoke habitually?’, and the imperative ‘Sam, smoke habitually!’ all share the same propositional content, expressed by the sentence-part ‘that Sam smokes habitually’. On Searle’s account of propositions, a speaker in uttering any one of these sentences performs the act of referring to Sam and the act of predicating the expression ‘smoking habitually’ of him. That is, the three different illocutionary acts all express the same proposition because the same propositional act is performed in each case. In each case, however, the proposition is expressed with a different *illocutionary force*: the first sentence has the illocutionary force of a statement, the second sentence has the illocutionary force of a question, and the third sentence has the illocutionary force of a command. The grammatical correlate of the illocutionary act is a complete sentence. In order to warn, describe, request, or question, etc., one needs to utter a complete sentence.<sup>3</sup>

The expression ‘I state’ in the sentence ‘I state that Sam smokes habitually’ is an explicit *illocutionary force indicator*, telling us directly what illocutionary act is being performed by the speaker in the utterance of the sentence. Searle notes that the following items can act as illocutionary force indicators in English: word order,

---

<sup>3</sup> Searle points out that this may be a one word sentence such as ‘Stop’ (1969: 25).

stress, intonation, punctuation, the mood of the verb, and, as in our example, performative verbs (1969, 30). It is important to recognize that *any* complete sentence will have by virtue of its meaning some kind of illocutionary force indicator which determines the type of illocutionary act that a literal and serious utterance of the sentence will constitute. Searle stresses that we should not take the primary distinction to be between sentence meaning and illocutionary force, that is, between a sentence and the illocutionary act performed in the utterance of the sentence.

Complete sentences, unlike propositions, are not neutral with regard to illocutionary force. The serious and literal utterance of a sentence therefore will always include its being uttered with a particular illocutionary force. Illocutionary force indicators of some sort invariably form part of the meaning of a sentence. The crucial distinction, Searle argues, is between *illocutionary force* and *propositional content*.

How do the details of Searle's theory help with the problem about moral motivation? Recall the internalist intuition that motivation is conceptually tied to moral judgement. If we think about judgements or statements in terms of their propositional content, that is, solely in terms of what they assert, then we will tend to view the internalist's conceptual claim as a claim about the conceptual connections between the propositional content of moral judgements and motivation. It will then be tempting to say that motivation is conceptually tied to the moral terms and phrases which distinguish the propositional content of moral judgements from non-moral judgements. And thus we might conclude that it is part of the truth-conditional meaning of terms such as 'right' and 'good' that an agent judging something as right or good is motivated to conform to her judgement.<sup>4</sup> However, Searle's distinction between the illocutionary force and the propositional content of an utterance suggests another way of interpreting the internalist intuition: namely, as a claim about the illocutionary force rather than the truth-conditional content of moral utterances.

---

<sup>4</sup> In chap 1, I give a number of reasons for rejecting this form of content internalism.

Speech act theory gives us a structured way of understanding the conditions governing the performance of illocutionary acts and explains how these acts relate to their propositional contents. Consider the simple case of assertions. To successfully assert that, 'Churchill was a wartime prime minister,' I must satisfy what Searle terms the 'essential condition.' For assertions the essential condition is that the utterance 'counts as an undertaking to the effect that p [e.g. That Churchill was a wartime prime minister] represents an actual state of affairs' (Searle 1969: 66). It is clear that whether or not this condition is satisfied, and thus whether or not I successfully perform the assertion, is independent of the truth of the proposition asserted. That is, the proposition, 'Churchill was a wartime prime minister,' is true or false regardless of whether I succeed in asserting it. We might say, then, that there is a necessary connection between the illocutionary act of asserting that p and the essential condition that the utterance involved counts as an undertaking to the effect that p represents an actual state of affairs. This is perfectly consistent with there being no necessary connection between the truth of the proposition asserted and the satisfaction of the performance condition in question. In a similar way, we might construe the motivational requirement on moral judgements as one of the conditions regulating the performance of moral utterances, rather than as a feature of the propositional content of such utterances. If I say, for example, 'Churchill's support for saturation bombing was morally wrong,' this can be true or false regardless of whether or not I successfully perform the moral speech act. Even so, we might judge my performance of the moral speech act as faulty, in a way yet to be specified, if I am not appropriately motivated.

In chapter one, I said that the expression 'moral judgement' is structurally ambiguous: it can refer either to the act of making a moral judgement or to the content of a moral judgement. In this section I have suggested that speech act theory provides a way of understanding this distinction in terms of the difference between the performance of a particular kind of linguistic act and the propositional content of such an act. While this is a helpful first step, we need to give a precise account of how the connection between illocutionary force and motivation will work in the case of moral judgements.

#### 4. Assertion and Compound Speech Acts

The last section showed that even in the case of a simple assertion we can distinguish between the conditions governing the performance of the illocutionary act of assertion and the truth conditions for the proposition asserted. And I noted that these two features of the speech act are logically independent. But does the example of an assertive speech act help us with the case of moral judgements? It is not clear that it does. Consider the conditions that Searle says are necessary for the successful and sincere performance of the illocutionary act of assertion:

- a) *Propositional content condition.* Any proposition p.
- b) *Preparatory conditions.*
  - 1. The speaker has evidence (reasons, etc.) for the truth of p.
  - 2. It is not obvious to the speaker that the hearer knows p.
- c) *Sincerity condition.* The speaker believes that p.
- d) *Essential condition.* The utterance counts as an undertaking to the effect that p represents an actual state of affairs. (Searle 1969: 66)

What this list tells us is that there is nothing in the conditions governing the successful and sincere performance of the illocutionary act of assertion that entails that the speaker holds a particular motivating attitude or desire towards the propositional content of the illocution. And this seems to count against the attempt to meet the internalist requirement on the basis of the conditions governing the performance of moral assertions. Does this mean that to accommodate internalism we must return to the notion that the motivational requirement is built into the meaning of moral terms or the truth conditions for moral judgements? I don't think so. We need to examine a slightly more complex case of assertion to see why the absence of a motivational requirement in the conditions on assertion does not undermine a speech act approach to moral motivation.

Consider the following example. In the context of a lively discussion with a friend, Emma issues the following utterance: 'I argue that Churchill was at one time a member of the Liberal Party.' This utterance clearly involves an act of assertion. Emma asserts that Churchill was once a member of the Liberal Party. But that is not all she does. The presence of the illocutionary verb 'to argue' tells us that she is trying to convince her friend of the truth of her assertion. As Searle points out, to say

‘I am arguing that p but not attempting to convince you that p’ sounds inconsistent. By contrast, saying ‘I am simply stating that p but not attempting to convince you that p’ seems acceptable (Searle 1969: 66). Something is going on in the case of arguing that p that is absent in the case of simply asserting that p. I suggest that in the case of arguing that p one is attempting to convince the hearer of the truth of p. In the vocabulary of speech act theory, we might say that arguing that p involves the speaker’s issuing a directive illocution along the lines of ‘believe that p!’ This is in addition to the act of asserting that p.

The verb ‘to argue’ is simply one example of an illocutionary verb which when used in the context of a speech situation involves the performance of more than one illocutionary act.<sup>5</sup> Another example of this kind of ‘compound’ illocutionary verb is the verb ‘to warn.’ If a speaker says, ‘I warn you that the bridge is about to collapse,’ she asserts that the bridge is about to collapse and directs the hearer to avoid the danger. That is, she performs both an assertive and a directive illocutionary act. If I say to you, ‘Let’s go to the movies tonight,’ I direct you to go to the movies and commit myself to going with you. So I perform a directive illocutionary act and an act of commitment.

What these examples show is that we need not worry that the performance conditions on bare assertion do not involve the speaker’s having a particular motivational attitude towards the asserted proposition. We can explain the presence of a motivational requirement on moral speech acts by establishing that their performance invariably involves the performance of illocutionary acts linked to motivating states. In the case of moral speech acts, we need to determine what kinds of illocutionary acts are standardly performed in the sincere and literal utterance of a moral judgement.

---

<sup>5</sup> Fotion discusses this aspect of Searle’s theory of speech acts. He refers to what I call compound speech acts as ‘double speech acts’ (2000: 57-64).



The fact that more than one illocutionary act is performed by way of a single utterance is not by itself sufficient to explain the internalist intuitions with which we are concerned. There are some expressions which are conventionally and systematically used to perform illocutionary acts other than those that make up their literal meanings.<sup>6</sup> For example, the question, ‘Can you pass the salt?’, is almost always used to make a request, even though, taken literally, it makes no request of the hearer to pass the salt. In a sense, we are already perfectly aware that moral utterances are conventionally used to do more than issue bare assertions. What we are after is an account of why the connection between the various illocutionary acts seems so close in the case of moral judgements.

Consider again the compound illocutionary verb, ‘to argue.’ In order to argue for a proposition, the speaker must assert its truth and direct the hearer to accept it. Arguing is a distinct kind of speech act comprising a number of basic illocutionary acts unified by a particular goal: that of *convincing* the hearer about the truth of the proposition presented. Is there a distinct kind of compound speech act involved in making moral judgements? It seems to me there is: moral judgements are a type of *evaluation*. Perhaps if we think of moral speech acts as evaluative illocutions rather than bare assertions we may move some way towards understanding the different illocutionary acts involved in moral judgement. Of course, moral judgements are only one type of evaluation; we also make, for example, epistemic, inferential, aesthetic, and prudential evaluations. While it is intuitively easy to distinguish moral evaluations from the first three types of evaluation, the shared practical orientation of prudential and moral evaluation makes distinguishing between them more difficult. However, this is not the place to undertake the difficult task of giving a precise account of how morality differs from other types of practical evaluation. For our purposes, it will be adequate to talk of moral evaluation as a form of practical evaluation: as one way in which we assess conduct in determining how to act.

---

<sup>6</sup> See chap 2 of Searle (1979) for a detailed analysis of the phenomenon of ‘indirect speech acts’.

In order to understand the nature of moral speech acts it will help to outline the conditions that might be said to govern the performance of moral evaluations. Before attempting this, however, I want to examine in more detail elements of Searle's theory of illocutionary acts.

## 5. The Nature of Illocutionary Acts

Searle's example of assertion shows that there are a number of different types of performance conditions he thinks illocutionary acts must satisfy in order to be successfully and sincerely performed. In this section, I discuss his classification of illocutionary acts based on differences in the way they satisfy these conditions.

According to Searle 'there are a...limited number of basic things we do with language: we tell people how things are, we try to get them to do things, we commit ourselves to doing things, we express feelings and attitudes, and we bring about changes through our utterances' (1979: 29). The most important notion with regard to classifying the uses of language is the essential point or purpose of a speech act (Searle 1969: 69; 1979: 2). Searle calls this the act's 'illocutionary point' and says that it corresponds to the notion of an essential condition (1979: 2).<sup>7</sup> For example, if we look at the essential condition for assertives, given above, it tells us that the illocutionary point of assertive speech acts is to represent the world as being a certain way. By contrast, says Searle, the essential point of a command is to get the hearer to do something, and the essential point of a promise is to place oneself under an obligation to perform a certain act. Using the basic notion of differences in illocutionary point, Searle classifies illocutionary acts into five basic types: i) *assertives* (e.g., statements, descriptions); ii) *directives* (e.g., requests, commands, questions); iii) *commissives* (e.g., promises, commitments); iii) *expressives* (e.g., congratulations, apologies); iv) *declarations* (e.g., appointing, christening) (1979: 12-20).

---

<sup>7</sup> The analysis in terms of performance 'conditions' comes from Searle's *Speech Acts* (1969). He introduces the notion of 'illocutionary point' in developing a detailed taxonomy of illocutionary acts in a later piece of work (Searle: 1979: chap. 1). Given that the later analysis supplements rather than replaces the earlier one, I draw from both in applying Searle's theory to the particular problem of moral judgements.

In addition to the notion of an essential or illocutionary point, Searle claims that the psychological state expressed in the performance of a speech act is crucial in identifying its type (Searle 1969: 60; 1979: 4). Thus, each main category of illocutionary act has a particular type of psychological state the expression of which is necessary for the sincere performance of acts within that class. The sincerity condition for assertives, shown above, is that the speaker believes what she asserts. By contrast, the sincerity condition for directives is that the speaker has a desire that the hearer do the act the speaker directs her to do. For example, if I ask you to shut the window, I am sincere in so far as I want or desire you to shut the window. The sincerity condition for commissives is that the speaker intends to perform the act she has said she will do: if I promise to meet you at the theatre, I am sincere in so far as I intend to meet you there.

Sincerity conditions are distinct from the other types of performance conditions in that they do not have to be satisfied in order for an illocutionary act to be successfully performed (Searle 1969: 62; 1979: 4-5). We thus distinguish between sincerity conditions and success conditions.<sup>8</sup> Although an agent can be insincere in her performance of an illocutionary act and fail to have the appropriate psychological state, in performing the act she still expresses or purports to have the psychological state required by the sincerity condition.<sup>9</sup> To deny giving this special status to

---

<sup>8</sup> If an illocutionary act, say a promise, is unsuccessfully performed it does not count as a promise. If the promise is merely insincere, the speaker succeeds in performing the promise even though the performance is defective. In Austin's terms, an unsuccessful attempt at an illocutionary act is a 'misfire', and a successful but insincere illocutionary act involves an 'abuse' of the relevant procedure (1962: chaps 3, 4).

<sup>9</sup> Mike Ridge argues that there are cases in which an agent fails to have the psychological state necessary to satisfy the sincerity condition for the illocutionary act she performs and yet it seems her speech act is sincere or at least not insincere (Ridge 2002; 2004). For example, Emma might believe falsely that she wants her friend to come with her to the cinema. When Emma requests that her friend accompany her to the cinema, on Searle's model, she is sincere insofar as she desires him to accompany her. But Ridge points out that in cases like these, given that Emma believes she wants her friend to come with her, it seems that Emma is sincere in her request. Ridge argues that to make sense of these kinds of examples we need to distinguish between acceptance conditions and sincerity conditions. Roughly, a speech act is sincere just in case the speaker has a second order belief that she possesses the psychological state it is the function of that kind of speech act to express (2004: 6). The first order states that such speech acts conventionally express are, on this account, the acceptance conditions for the speech acts. This leads Ridge to formulate what he takes to be the most plausible version of moral judgement internalism. Acceptance judgement internalism posits a necessary connection between the acceptance of a moral judgement and the possession of a relevant motivating state (Ridge 2002: 7). Acceptance is thus a psychological rather than a linguistic constraint on moral judgement. Ridge argues that the real issue of moral motivation is thus the connection between

sincerity conditions would raise a number of difficulties. For example, in the case of promises, denying this status would mean that we could not hold someone accountable for their promise if they were insincere, as they would not have made a promise. More generally, we could not easily speak of insincere statements or insincere requests, as such illocutions would either be sincere or not made.

The essential condition and the sincerity condition for a particular illocutionary act allow us to determine its direction of fit. This is a notion familiar from our discussion of the desire/belief thesis in chapter one. Searle applies it to the way in which the propositional content of an illocutionary act relates to the world (1979: 3). For example, in making an assertion, a speaker attempts to get her words (the propositional content) to fit the world, whereas in issuing a directive, a speaker attempts to get the world to conform to her words (propositional content). An assertion is thus said to have a word-to-world direction of fit while a directive is said to have a world-to-word direction of fit. Commissives, like directives, have a world-to-word direction of fit: if I commit myself to meeting you tomorrow, I aim to get my behaviour (the world) to fit what I have said I will do (my words).

Expressives and declarations are slightly more complex than the other forms of illocutionary acts. The essential point of an expressive act is to express the psychological state required by its sincerity condition. The precise nature of this psychological state depends on the particular type of expressive illocutionary act being performed. For example, if I congratulate you on winning the competition, I am sincere in so far as I am pleased for you and respect your achievement. By contrast, if I apologise for not posting your entry form, I am sincere in so far as I feel sorry for having forgotten. Expressives do not have a direction of fit: they are neither attempts to get the world to fit the words as with directives nor to get the words to fit with the world as with assertives. Declarations, by contrast, do not have a sincerity condition but have a dual direction of fit. The essential point of a

---

*thinking* that something is right and motivation not between saying that something is right and motivation. In chap 5, I argue that the linguistic rules governing moral speech acts are an integral part of moral practice, and that moral thought, insofar as it reflects *that* practice, is also governed by the same rules.

declaration is to bring about a match between the propositional content of an utterance and reality. In performing a declaration we aim both to get the words to fit the world and to get the world to fit the words. For example, if I declare the meeting adjourned, I attempt to get the words to match the world by making it the case, by virtue of my very utterance, that the meeting is adjourned: by making the world match the words.

This short overview of Searle's theory of illocutionary acts provides us with the background needed to attempt a more detailed analysis of moral speech acts. The approach I favour aims to elucidate the nature of moral judgements by determining how moral illocutions meet the various kinds of performance conditions outlined above.

## **6. How to Make a Moral Judgement**

In this section, I use Searle's model, given above for assertion, to provide an account of the conditions governing the performance of moral evaluations.<sup>10</sup> The following conditions, then, are necessary for the sincere and successful performance of the illocutionary act of moral evaluation.<sup>11</sup>

### *i) Propositional content condition*

In the case of evaluations, unlike bare assertions, there is a feature of the proposition expressed that acts as an illocutionary force indicating device: that tells us that an evaluation rather than a bare assertion or some other illocution is being performed. The feature in question is the evaluative predicate or evaluative term. In using evaluative terms such as 'good', 'right', or 'justified', a speaker signals to her audience that she is making an evaluative judgement. Let's say that a condition on the propositional content of evaluations is that the speaker predicates an evaluative

---

<sup>10</sup> For ease of expression and depending on the context, I will sometimes use 'practical evaluation' or 'evaluation' rather than 'moral evaluation'.

<sup>11</sup> In order to elucidate the general structure of illocutionary acts, Searle devotes an entire chapter to outlining the necessary and *sufficient* conditions for the successful and sincere performance of the illocutionary act of promising (1969: chap. 3). Here I attempt the more limited task of outlining the conditions necessary for the successful and sincere performance of the illocutionary act of moral evaluation. I make no claim that the analysis provides conditions which when taken together are sufficient for performing successful moral speech acts.

property of the object referred to by the subject term of the uttered sentence. For example, in the moral evaluation, 'Supporting the strike is the right thing to do,' the property of rightness is predicated of the act of supporting the strike.

*ii) Essential Condition*

I have argued that evaluations involve something *in addition* to assertion, rather than their being distinct from assertions, in the way, for example, that imperatives are distinct from assertions. An evaluation is in part made up of an assertion, so the conditions governing assertion also apply to evaluations. This means that the essential condition for assertion must be incorporated into our statement of the essential condition for evaluation. What then is the essential point or aim of moral evaluations? That is, what is it that the speaker wants to accomplish by issuing an evaluative speech act? I suggest we think of practically evaluative utterances as essentially involving a commitment by the speaker to act in accordance with the content of the uttered proposition: that is, a commitment to being appropriately guided by those features of a situation the utterance represents as favourable or unfavourable. Moral evaluations thus involve, in addition to assertion, the performance of a commissive illocutionary act: an act that commits the speaker to behaving in a certain way in the future.

On this model, an evaluative utterance counts as an attempt to represent an actual state of affairs and so satisfies the essential condition for assertion. However, the point of an evaluative utterance is not just to represent any old state of affairs. Rather, and in line with the propositional content condition already discussed, its aim is to represent those features of a situation that are of practical normative significance: that favour acting in one way rather than another. The fact that the utterance also constitutes a commitment to be guided by the content of the evaluation is integral to its being practical in both its subject matter and use. Only if an agent is committed to being guided by her practical evaluations can she be said to be engaging in distinctly practical deliberation.



### *iii) Preparatory Condition*

Preparatory conditions involve the kinds of things that need to be presupposed before a speaker can begin legitimately to issue her speech act. In general they concern the suitability of a speech situation for the satisfaction of the point of a speech act as described by its essential condition. Searle explains:

For example, if I make a request to someone to do something which it is obvious that he is already doing or is about to do quite independently of the request, then my request is pointless and to that extent defective. In an actual speech situation, listeners, knowing the rules for performing illocutionary acts, will assume that this condition is satisfied. (Searle 1969:59)

The relation between evaluation and assertion means that the preparatory conditions for assertion also apply to evaluation, so the speaker must have evidence for the truth of her evaluative proposition. A preparatory condition specific to evaluative illocutions is that the speaker issuing the illocution must be *normatively engaged*: that is, the speaker must be able to be moved by normative considerations.<sup>12</sup> A practical evaluation is defective, then, if the speaker is unable to be influenced by and to respond appropriately to practically normative considerations. This is because if a speaker is outside practical normativity in this sense, then her moral evaluation is pointless, given that its aim is to guide her action. And this means that a hearer or hearers will assume that the speaker issuing the evaluation is able to be guided by her judgement.

### *iv) Sincerity Condition*

This condition looks like it might provide the best opportunity to explain the close connection between evaluative judgements and motivation. After all, motivating states are mental states, and the sincerity condition for illocutionary acts involves the mental state or states that an agent must possess in order for her illocution to be sincere. Can we say, then, that the possession of a motivating attitude of some sort is necessary for the sincere utterance of a practical evaluation? This will depend on which illocutionary acts we take to make up the 'compound' act of evaluation. The

---

<sup>12</sup> This requirement is admittedly vague at the moment. A more precise account of it will be given in the context of a discussion of the notion of 'normative expectation' in chap. 4.

essential condition tells us that an evaluative speech act consists of an assertive illocutionary act and a commissive illocutionary act. The assertive portion represents the evaluative proposition as true and the commissive portion commits the speaker to being guided by the content of the proposition. As we have seen, making an assertion does not entail the possession of a motivating attitude; consequently, to elucidate the link between evaluation and motivation we must focus on the commissive act.

The essential point of commissives is to commit the speaker to some future act or type of act. The archetypal commissive is promising. If I promise to do A, then I commit myself to doing A. Promising involves, essentially, the speaker's placing herself under an obligation to act. The sincerity condition for commissives also connects them directly with action. To commit myself sincerely to acting in a certain way, I must intend to act in such a way. For example, to commit myself to being guided by an evaluative proposition, I must *intend* to be guided by that proposition. And intention is a motivating attitude *par excellence*.<sup>13</sup>

To summarise, then, I have argued that the following conditions are necessary (though not sufficient) for the successful and sincere performance of the speech act of moral evaluation:

- a) *Propositional content condition*. The speaker predicates an evaluative property of the object referred to by the subject term of the uttered sentence.
- b) *Preparatory conditions*.
  1. The speaker has evidence (reasons, etc.) for the truth of p (viz: the evaluative proposition).
  2. It is not obvious to the speaker that the hearer knows p.
  3. The speaker is normatively engaged.

---

<sup>13</sup> It will become clear when we discuss the issue of failures of motivation in the next chapter that this formulation of the sincerity condition needs to be refined. However, it is a useful first step and points towards the formulation for which I finally opt.

c) *Sincerity conditions.*

1. The speaker believes that p.
2. The speaker intends to be guided by the content of p.

d) *Essential condition.*

1. The utterance counts as an undertaking by the speaker to be guided by the content of p.
2. The utterance counts as an undertaking to the effect that p represents an actual state of affairs (viz: the features of the situation that favour acting in a particular way).

Thus far in the chapter, I have given an account of act internalism in terms of the performance of a certain kind of speech act called a moral evaluation. I have argued that we can accommodate the fact-stating and the motivating aspects of moral judgements by understanding them as a type of compound speech act. According to this approach, a moral evaluation involves the ‘compound’ performance of both an assertive illocutionary act and a commissive illocutionary act. The connection between moral judgements and motivation is secured by the essential point of such evaluations, namely, that a moral utterance counts as an undertaking by a speaker to be guided by the content of the uttered evaluative proposition.

On the speech act model, a speaker in making a moral judgement expresses her commitment to acting in accordance with the content of the uttered proposition. In making a *sincere* moral judgement, the speaker must intend to be guided by the content of the uttered proposition. Intention is a motivating attitude *par excellence*. Consequently, an agent who makes a sincere moral judgement will have some motivation to act in accordance with it. In so far as the speaker does not understand that her utterance commits her to acting in accordance with the evaluation, her utterance will not count as a moral judgement. More generally, if a speaker fails to satisfy the essential condition of a speech act, the speech act will not count as having been successfully performed. Nonetheless, because the motivational requirement on moral judgement is accommodated at the level of illocutionary force rather than at the level of propositional content, the speech act approach does not threaten the

category of moral requirements. That is, moral claims can be true or false independently of whether the agent who makes them is motivated to conform to them.

In the next section, I discuss an approach to the issues of internalism and moral motivation by David Copp (2001). Copp's approach is similar to mine in that he attempts to accommodate moral judgement internalism and moral cognitivism by construing the internalist requirement in terms of the linguistic conventions governing the performance of moral judgements. I argue that there are a number of reasons for preferring my account to his.

### **7. Copp's Discourse Internalism**

In a recent paper, David Copp proposes addressing the problem of moral motivation from the perspective of language (2001). His suggestion is that the connection between moral judgement and motivation be explicated via the linguistic conventions governing the performance of moral speech acts, rather than at the level of truth-conditional content. Copp calls his position 'realist-expressivism.' He argues, as I did in chapter one, that the traditional opposition between moral realism and antirealist-expressivism does not derive from the expressivist portion of the antirealist position. That is, it does not follow from the fact that moral judgements typically express affective or conative states that some form of non-cognitivist moral anti-realism is true. Expressivism as a positive thesis about the linguistic function of certain kinds of evaluative sentences is perfectly consistent with cognitivism about the content of those sentences, and thus with realism about the evaluative facts and properties to which such sentences refer.

Whereas I apply Searle's work on speech acts to the problem of moral motivation, Copp draws on Grice's notions of conversational and conventional implicature as well as Frege's concept of 'colouring' to develop his view. This means that while the general thrust of our positions is the same – a linguistic approach to the problem of moral motivation – the details are quite different. Copp says that his view involves the pragmatics of moral *assertion*. He argues that it is in virtue of the

conditions on the performance of moral assertions that speaker's will standardly evince a conative attitude of the relevant sort (Copp 2001: 14). And this in turn is explained by the semantics of moral terms: it is because of the non-truth-conditional meaning or 'colouring' of moral terms that moral assertions conventionally express conative attitudes. Consequently, says Copp, the 'having of the relevant belief and the having of the appropriate motivation might each be necessary to the sincerity of the speaker's assertion without the motivation being a necessary condition of belief' (Copp 2001: 38).

As we have seen, our experience of morality suggests that there are both cognitive and affective components to genuine moral judgement. Copp needs to explain why his proposal is not merely an *ad hoc* restatement of this fact. His solution is to use the Fregean notion of colouring to explain how some predicates have a 'broad' or non-truth-conditional aspect to their meaning, in addition to their 'core' or truth-conditional meaning, which affects their performance conditions when used to make assertions (Copp 2001: 15). Copp suggests that pejorative terms give a clear example of how such colouring works. He uses the pejorative term 'cur' to illustrate the point. While 'cur' and 'mongrel dog' have the same reference, the colouring of 'cur' means that it is standardly also used to express contempt for a dog. Thus, while, 'The mongrel dog howled all night,' and, 'The cur howled all night,' have the same reference and express the same thought, the colouring of 'cur' means that the utterance of the latter expression also implies contempt for the dog. I will use a different example in order to avoid difficulties arising from the fact that 'mongrel' is also used as a term of abuse. What we require is an instance of a clearly coloured term and its non-coloured equivalent to elucidate the way in which the colouring of a term can explain the reliable and conceptual connections between its use and, for example, expressions of contempt.

The term 'pinko' is a derogatory and offensive word used to describe people who hold left-wing or socialist political views. The term is used to refer to such people but also to express the speaker's contempt for them. If this is right, then it would indicate a lack of mastery of the conventions for the use of the term 'pinko' if one

were to use it simply to refer to socialists without also to expressing contempt for them. Even so, an assertion such as, 'The teacher is a pinko,' can be true even if the speech performance involved is defective because the speaker does not express contempt. The proposition believed is the same in the case of this statement and its non-coloured equivalent, 'The teacher is a socialist,' as are the truth-conditions for both statements. What differs is the contempt implied by judgements in which the term 'pinko' is used. The reliable connection between such judgements and contempt is a result of the linguistic conventions governing the use of the term 'pinko' rather than the truth-conditional content of such judgements, or an external and contingent desire to condemn 'pinkos.' Copp calls this kind of internalism, 'discourse internalism'.

Copp holds that the use of coloured predicates provides an example of grammatically assertive utterances that require the expression of a relevant conative state as well as the usual cognitive state in order to be linguistically appropriate. He applies this model to the use of moral terms, highlighting the difference in sincerity conditions between assertions such as, 'Cursing is widespread,' and those such as, 'Cursing is morally wrong' (2001: 14). The speaker of the first assertion is sincere if she has a relevant belief. The speaker of the second assertion is sincere if she has a relevant belief *and* is in a relevant conative state. Copp argues that it is the colouring of the expression 'morally wrong' that accounts for the difference in sincerity conditions between the two assertions. And he holds that this colouring is part of the meaning of 'morally wrong,' 'at least in a wide sense of 'meaning,' *for it is characteristic of the linguistic conventions governing the use of the term*' (Copp 2001: 14. My emphasis).

A problem with Copp's notion of wide meaning is that the presence of linguistic norms determining what a term or expression conventionally implies does not show that the term or expression in question has this 'coloured' aspect as part of its literal meaning. For example, conventions governing the use of the sentence, 'You are blocking my view,' dictate that it is standardly uttered as a request for the hearer to get out of the speaker's line of vision. A sincere use of the sentence would normally,



therefore, require the speaker both to believe that the hearer was blocking his view and to desire the hearer to move. However, this does not show that the literal meaning of the sentence includes a hidden or coloured directive aspect. Rather, the literal and non-coloured meaning of the sentence makes it *appropriate* for issuing indirect requests. A speaker is insincere if she uses this sentence in a conventional way to issue a request but does not desire the hearer to move. However, the sentence can be uttered sincerely and literally as a direct assertion without a request being made. For example, I might request that my friend sit in front of me to obscure the view of a frightening movie. When she is in the correct position I say, 'You are blocking my view,' and mean just what I say. The sincerity of my assertion in such cases depends only on my possession of the relevant belief. None of this alters the fact that someone who did not understand that such sentences are conventionally used to make requests would lack mastery of the linguistic conventions regulating their use. These considerations suggest that it cannot merely be the presence of linguistic norms governing the standard uses of terms and expressions that makes those uses part of literal meaning.

The theory of indirect speech acts shows that we can generally isolate the literal meaning of a sentence from its indirect meaning, even when the indirect meaning forms part of the conventional use of the sentence (Searle 1979: chap. 2). However, there may be some terms whose conditions of use cannot be separated from their literal meanings. If there were predicates that not only referred to particular properties but invariably included expressive force indication as part of their literal meaning, then they might provide an example of colouring not explicable in terms of indirect speech acts. Pejoratives are characterised by the role they play in the performance of expressive illocutionary acts, specifically in the expression of contempt. For example, the literal and sincere assertion that someone is a 'nigger' invariably includes an expression of contempt for that person. And it is difficult to imagine a context in which the non-ironic assertive use of 'nigger' would not convey contempt. However, even if there were invariably coloured predicates, it is doubtful that it would be in virtue of the conditions on assertion that a speaker in using them expressed, for example, contempt. Rather, the colouring of the predicate would

function as a separate force indicator for the relevant non-assertoric illocutionary act which would be performed *in addition* to the assertive act.

There are a number of considerations that support the claim that it is not in virtue of the conditions on the performance of assertions that the use of coloured predicates involves the expression of conative states. As its name suggests, Copp's realist-expressivism is an attempt to explain moral motivation along expressivist lines. An expressivist account of pejoratives would hold that a term such as 'pinko' is not a genuine referring term. In so far as it refers at all, it refers to the property of being a person with left-wing or socialist political views. We can factor out pejorative terms into a genuine predicate and an expressive force indicator. Thus, 'pinko' might be represented as, B! (socialist), where 'B!' indicates an expression of contempt.<sup>14</sup> The sentence, 'The teacher is a B! (socialist)' would therefore be equivalent to the sentence, 'The teacher is a pinko.' This suggests that it is not by way of the conditions governing the assertive portion of pejorative sentences that speakers express contempt. The reason is that it is not possible to *assert* that someone is a 'pinko,' only that someone is a socialist. But this means that it cannot be in virtue of the conditions governing the *assertion* of such terms that they standardly express conative states of the relevant kind.

The nature of the performance conditions governing assertive illocutionary acts also lends support to the claim that it is not by way of making assertions that speakers express conative states. Assertives are said to have a 'word-to-world' direction of fit similar to the mind-to-world direction of fit of cognitive states. The illocutionary point of assertives is to say that things stand in a certain way and assertives commit speakers to the truth of the propositions asserted. The assertion of a statement is thus the standard linguistic device for the expression of belief. It is commonly held that it is not part of the nature of belief-like states to motivate agents. It seems reasonable to assume that, equally, it is not the task of assertives to express speakers' non-

---

<sup>14</sup> Blackburn uses this style of 'expressive' language in order to give an analysis of standard indicative moral sentences in expressivist terms: 'Imagine a language unlike English in containing no evaluative predicates...It might contain a 'hooray!' operator and a 'boo!' operator (H!, B!) which attach to descriptions of things to result in expressions of attitude' (Blackburn 1984: 193).

cognitive states. If this is correct, the role of expressing conative states should not be attributed to the assertive component of speech acts.

If the above arguments are sound, we have reason to reject Copp's account of colouring, based as it is on a flawed account of how pejoratives work. What implications does this have for his attempt to view the expressive force of moral predicates along the lines of pejoratives? In so far as moral predicates do function like pejoratives, the expressive force of moral judgements will not be explicable via the conditions on the performance of assertive speech acts. However, we might question the extent to which moral predicates do work like pejoratives. Unlike moral terms, pejoratives appear to have expressive force regardless of the grammatical context in which they are used (cf. Searle 1969: 156). For example, if I ask the question, 'Is the teacher a pinko?', I express contempt for socialists even though I make no assertion. Similarly, if I issue the command, 'Catch the pinko!', I express contempt for socialists.

Now, the neutrality of predicates with regard to force is crucial in distinguishing the propositional content of a speech act from its illocutionary force. The illocutionary force of a speech act determines how the propositional content relates to the world, that is, whether the proposition is to be understood as an assertion, a request, an expression of feeling or a declaration, and so on. Pejoratives, at least as understood by Copp, are not neutral in this way and do not refer to a property distinct from that referred to by their non-pejorative equivalents. Thus, whereas the proposition represented by the sentence-part, 'That the teacher is a socialist,' is force-neutral, the proposition represented by the sentence-part, 'That the teacher is a pinko,' is not: it includes expressive illocutionary force indication, in this case, contempt for socialists. But this suggests that the latter sentence-part cannot play the role a genuine proposition must play on the speech act model.

In the case of pejoratives, it is relatively easy to isolate genuine predicates from force indicators. If we take the sentence, 'The teacher is a pinko', it clearly has assertive and expressive force indicators as part of its meaning, determining how the propositional content, 'That the teacher is a socialist', relates to the world. 'Socialist' rather than 'pinko' is thus the genuine force-neutral predicate. But there seems to be no comparable way to factor thin evaluative predicates into a genuine referring component and a force-indicating component. Unlike anti-realist expressivists, Copp holds that moral predicates refer to genuine evaluative properties as well functioning as expressive force indicators. Consequently, it is not an option for him to factor an evaluative judgement into the non-evaluative properties genuinely referred to by the sentence and the expressive force indication given by the evaluative terms used in the sentence. For example, he cannot interpret, 'Teachers are morally good,' as 'H! (teachers)'. However, this makes it difficult to see how evaluative predicates can play the force-neutral role that genuine predicates play in the propositional component of speech acts.

The problem with tying expressive force to the literal, if non-truth conditional, meaning of a predicate is related to the issue of unasserted contexts. The claim that a predicate must be able to occur with the same meaning in unasserted as well as asserted contexts has been used to undermine traditional non-cognitivist or expressivist theories of the meaning of moral terms.<sup>15</sup> Any adequate theory of meaning must account for the fact that predicates can have the same meaning whether they occur in assertions, questions, conditionals, negations, and so on (Searle 1969: 137). For example, if the assertion, 'It is right', is to function as an answer to the question, 'Is it right?', 'right' must have the same meaning in both contexts. But while an assertion of rightness plausibly involves an act of commendation, this does not seem to be the case with a question about rightness. The literal occurrence of 'right' can occur, that is, in the sincere and literal utterance of interrogative sentences without the speaker thereby commending or endorsing anything. And, as Searle points out, there seem to be an 'indefinite number' of

---

<sup>15</sup> See Geach (1965), and Searle (1969: chap. 6) for different versions of the argument that non-cognitivism is unable to account for the meaning of moral terms in unasserted contexts.

counter-examples of this kind to the thesis that the meaning of evaluative terms is exhausted by their commendatory or expressive function in assertive contexts (1969: 139).

An interesting feature of pejorative terms is that they retain their expressive force even when they occur in unasserted contexts. The question, 'Is he a pinko?', and the assertion, 'He is a pinko', both express contempt for socialists. This is one important disanalogy between pejorative terms and moral terms: moral terms do not have commendatory force in unasserted contexts. And this is precisely what is said to undermine expressivist theories of moral meaning. However, this is a positive outcome if we are concerned to defend the status of moral predicates as genuine predicates. The reason is that if moral predicates do not retain their commendatory force in unasserted contexts, they can play the force neutral role that genuine predicates play in the propositional component of speech acts.<sup>16</sup>

While I agree with the general thrust of Copp's discourse internalism, there a number of reasons to prefer the speech act approach I advocate. Copp argues that the motivational requirement on moral judgements holds via the conditions on moral assertion. However, the standard conditions governing assertive illocutions are ill-suited to accomplishing this task. Even if we view pejorative terms as having literal non truth conditional expressive meaning, it seems it is not by way of the conditions on assertion that the use of a pejorative requires an expression of contempt. Recall, my proposal is that moral judgements do not entail motivation *qua* their status as assertives; rather, the motivational requirement stems from the fact that in making a moral judgement speakers perform the illocutionary act of evaluation, and evaluations have different performance conditions to bare assertions: conditions that include some kind of pressure on action. There are also some important disanalogies between moral terms and pejorative terms that appear to undermine Copp's approach. Unlike moral terms, pejorative terms are not force-neutral, retaining their expressive force in different grammatical contexts. This means that unlike moral

---

<sup>16</sup> This assumes, of course, that moral predicates are neutral with regard to other kinds of illocutionary force. For example, it assumes that they are not essentially assertoric.

predicates, pejoratives cannot play the force-neutral role that genuine predicates play in the propositional component of speech acts. And this suggests that we should avoid an analogy with pejoratives if we are concerned to defend the status of moral predicates as genuine.

## 8. Evaluative Predicates and Force Indication

The discussion so far suggests that there is nothing unusual or queer about the apparent dual direction of fit of moral judgements, at least from the perspective of language. The difficult task is to give a detailed account of how in the case of moral judgements a single utterance involves the performance of more than one illocutionary act. I suggested we use the concept of a compound speech act to accomplish this task. In the above discussion we also came across another way in which a single utterance can be used to perform more than one illocutionary act: the indirect speech act. To begin this section I want briefly to review these two kinds of multiple speech acts.

As we have seen, *compound speech acts* involve illocutionary verbs or verb phrases which when uttered literally and sincerely in the context of a complete sentence allow the speaker to perform directly more than one illocutionary act. For example, 'I protest your treating the workers as inferior,' involves assertive, expressive, and directive speech acts. It asserts that there is something wrong with how the workers are being treated, it expresses disapproval of this, and it directs those responsible to do something to improve the situation. *Indirect speech acts* occur when a speaker utters a sentence with a particular illocutionary force, performs that illocution, but also performs another illocution that is *not* part of the literal meaning of the sentence (Searle 1979). In such cases there is a distinction between the literal meaning of the sentence and the intended force with which the sentence is uttered. For example, if I say, 'It is raining', and intend my utterance to be taken purely as a statement of fact, the illocutionary act of stating is performed directly using the conventional linguistic device (the indicative sentence) for making statements. However, if I utter the same sentence for the purpose of getting you to get your umbrella, the literal meaning of



the sentence and the intended force with which it is uttered come apart. The illocutionary act of requesting that you get your umbrella is performed indirectly by way of performing directly the act of stating that it is raining.<sup>17</sup>

The reason I did not construe the relation between the assertive and motivating aspects of moral judgements in terms of indirect speech acts is that such a model cannot adequately accommodate moral judgement internalism. This is because in those cases where the performance of a particular illocutionary act is only indirectly related to the direct performance of another act, it is always possible to perform the direct act on its own, even if doing so is linguistically unusual. For example, it is unusual but possible to issue purely as an assertive the sentence, 'You are standing on my foot'. Normally the utterance of such a sentence involves an indirect request for the hearer to get off the speaker's foot. By contrast, understanding moral judgements in terms of compound speech acts means that an agent making a moral judgement performs *directly* the illocutionary act of moral evaluation: the act sanctioned by the force indicators that are part of the literal meaning of moral sentences. On this model, the sincere and literal utterance of a moral sentence in the indicative mood is necessarily an evaluation rather than merely an assertion.

There is an apparent difference between most types of compound speech acts and the speech act of moral evaluation that raises some difficulties for the approach I advocate. In order to perform most compound speech acts *directly* one has to utter the relevant illocutionary verb. For example, to issue a protest directly, I must say, '*I protest* that such and such', and to issue a warning directly, I must say, '*I warn* you that such and such'. This is not the case with the majority of speech acts involving a single illocutionary act. And it doesn't seem to be the case with moral speech acts. If I want to make a statement directly, I don't need to prefix my utterance with 'I state'. Similarly, if I want to make a moral judgement directly I don't need to say, '*I evaluate* such and such as right'. In the case of single illocutionary acts this is

---

<sup>17</sup> Note that when one illocutionary act is performed indirectly via another, the direct illocutionary act is not replaced by the indirect act. Even though I request that you get the umbrella, my request is made by way of stating that it is raining. In fact, I would be unlikely to bring off the performance of the indirect act if my performance of the direct illocutionary act were defective.

explained by the fact that they generally have particular linguistic devices with force indicators sufficient for their direct performance. For example, the indicative sentence is the conventional device for performing assertives, and the imperative sentence is the conventional device for performing directives.

According to the compound approach to moral speech acts it can't be the case that a speaker says something is morally good without thereby performing a moral evaluation. But as with other compound speech acts, agents rarely perform moral speech acts using explicit illocutionary verbs.<sup>18</sup> They tend to say, 'Voting is good', rather than, 'I evaluate voting as good'. What needs to be explained is how moral speech acts can be performed directly without the use of explicit illocutionary verbs and why moral speech acts are distinct from most other forms of compound speech acts in being performable in this way. To understand the first issue we need to look at the force indicating role of evaluative predicates. If a speaker says, 'Giving the money back is right', the presence of the evaluative predicate 'right' in conjunction with other relevant force indicators tells us she is performing an evaluative illocutionary act. The assertoric force indicators are crucial here, because, as we have seen, evaluations involve assertions. This is why if a speaker asks, 'Is giving the money back right?', she cannot be performing directly an evaluative speech act. *A single predicate cannot by itself determine the force potential* of a sentence; rather, illocutionary force as part of the literal meaning of a sentence is determined by the interplay of a range of force indicating devices, such as word order, punctuation, the mood of the verb, performative verbs, and, I want to say, certain predicates. In the case of moral speech acts, these force indicators together determine the type of act constituted by the serious and literal utterance of a moral sentence.

---

<sup>18</sup> In practice agents rarely prefix their compound illocutionary acts with explicit force indicating verbs. Instead of saying, 'I protest your leaving work early every day', an agent might protest by making an assertion, e.g., 'You leave work early everyday', or by asking a question, e.g., 'Why do you leave work early everyday?'. In neither of these cases is the illocutionary force of the protest part of literal meaning of the sentence. So in both cases the protest is performed indirectly via the direct acts sanctioned by the force indicators that are part of the literal meaning of the sentences. And this means that it is perfectly possible for an agent to assert that someone is late for work everyday without protesting their lateness. This distinguishes such speech acts from moral speech acts which, I argue, cannot be performed without performing an evaluation, even in the absence of the relevant illocutionary verb.

What explains the fact that moral evaluations are unlike other compound speech acts in being directly performable in the absence of illocutionary verbs? I suspect one reason for this is the ubiquity of such evaluations. Human agents make evaluations, including moral evaluations, in a wide range of different circumstances and most of the time. By contrast, the contexts in which we need to make protests or issue warnings are relatively circumscribed. It makes sense, then, simply in terms of linguistic efficiency not to have to worry about uttering evaluative illocutionary verbs. In this sense evaluations are like assertions, imperatives and questions, but unlike most compound speech acts and commissives.<sup>19</sup>

A crucial distinction between the approach I advocate and Copp's theory is that on my account the motivating force of moral predicates is not part of their literal meaning, in a 'narrow' or a 'wide' sense of meaning. Given my position, however, I need to explain the fact that the connection between evaluative predicates and the act of evaluation is much closer than the connection between most other kinds of predicates and their roles in force indication. There does seem to be some kind of conceptual connection between evaluative predicates and the motivating attitudes typical of evaluation. This is why it is tempting to explain the meaning of evaluative predicates in terms of their *use* in commending or expressing approval. However, I argued above that we need to avoid doing this in order to preserve the force neutrality that enables moral predicates to function adequately in the propositional component of speech acts. To address these issues, I turn to a suggestion of Searle about how to understand the relation between evaluative terms and evaluative speech acts.

---

<sup>19</sup> To perform a commissive illocutionary act directly, the sentence uttered must, it seems, include a relevant illocutionary verb. If I say to you, 'I will meet you there tomorrow', I need not be committing myself to being there to meet you. The literal force of the sentence is that of a predictive assertion. For example, I may currently believe that I will be at the meeting tomorrow and so believe that I will meet you there tomorrow. This need involve no commitment on my behalf to be there for *you* tomorrow. By contrast, if I say literally and sincerely, '*I promise* to meet you there tomorrow,' I directly perform a commissive illocutionary act: I place myself under an obligation to be there to meet you tomorrow.

Searle agrees that we need to give some account of the apparent conceptual connection between evaluative predicates and the illocutionary acts typically engaged in when we use such predicates:

The data that we have to explain...are of these sorts: Calling something 'good' is characteristically praising, or commending, or recommending, or expressing approval of things so called. Furthermore, this seems not to be just a contingent fact, as is shown by the fact that the word 'good' itself is sometimes described as a term of praise... How can it be the case in these and other instances both that calling something W is indeed performing a speech act A, and yet it does not explain the meaning of W to say W is used to perform act A? (Searle 1969: 150)

The answer to this problem, suggests Searle, is to be found in the way that evaluative predicates are embedded in the institutions of assessment, appraisal, and evaluation in general (1969: 152). The point of evaluations or assessments is to rank or grade options: to determine which options are better, more favourable, of a higher standard, and so on, than others. And in order to do this an agent must assign an option a rank or grade in a relevant scale of assessment (Searle 1969: 151-152). Options can be given higher or lower evaluations, and we use evaluative predicates or 'grading terms' to indicate the level of assessment being given. In giving a high evaluation of something we typically use terms such as 'good', 'excellent', 'commendable', 'praiseworthy', and so on. And in giving a low evaluation of something we use, among others, terms such as 'poor', 'bad', and 'unsatisfactory'. So to call something 'good' or 'excellent' is necessarily to evaluate or assess it. Specifically, it is to give a relatively high evaluation or assessment of the item in question. And giving a high evaluation or assessment of something is typically to endorse or praise it.

Searle stresses that it does not follow from the fact that terms such as 'good' and 'excellent' are *used* to make high evaluations that they *mean* 'evaluated highly' (1969: 139, 152). To infer from the fact that a word is used to perform a certain illocutionary act that the meaning of the word is explicable in terms of that act is to commit the 'speech act fallacy' (Searle 1969: 136-141).

As an illustration of why this type of inference is problematic, consider the example of a simple case of assessment, in which a student's paper is marked as 'excellent'. An assessment of this kind entails that the marker evaluated the paper highly. But it does not entail that the paper is of an excellent standard. If I ask, 'Is the paper really of such a high standard?', I am not asking whether the paper was given a high evaluation by the marker; rather, I am asking whether the paper *deserves* or *warrants* such a high mark. Of course, I assume that in so far as the marker is sincere, she gave the high mark on the basis of her belief that the paper deserved or warranted such a mark. This simply underscores the fact that what she meant by calling the paper 'excellent' was not that it was evaluated highly by her, but that it deserved or warranted a high evaluation: that it satisfied the requirements for being given a high ranking on the relevant scale of assessment.

Such considerations indicate that the conceptual connections between evaluative predicates and acts of evaluation or assessment are explicable in terms of the way the predicates are embedded in the institutions of evaluation and assessment. If I evaluate something, I must assign it a grade or rank on the relevant scale of assessment. The standard device that allows me to accomplish this task is the evaluative predicate.<sup>20</sup> So evaluating something commits me to assigning it a rank or grade that can be expressed using an evaluative predicate. For example, in order to give an assessment of an essay, I cannot simply say, 'I evaluate the essay'; rather, I must give a more or less positive or more or less negative assessment of the essay (cf. Searle 1969: 51). Conversely, in giving either a positive or negative assessment of the essay, I am necessarily evaluating it. These points are captured by the following biconditional:

A particular act is an act of evaluation if and only if it involves an assignment of rank expressible by an evaluative predicate.

---

<sup>20</sup> I may not always be explicit in assigning a rank by way of a particular evaluative predicate. For example, I might say, 'I rate it very highly'. This seems to me to be an example of an indirect speech act: I perform an evaluation indirectly by asserting directly that I rate the option highly. This works partly because my rating it highly commits me to assigning it a rank expressible by a positive evaluative predicate. The hearer will typically, therefore, interpret my utterance as some form of positive evaluation.

It is crucial, however, to remember that such connections do not give the meaning of evaluative predicates, they simply indicate how we use them.

## **9. Review of the Argument So Far and Potential Difficulties for the Speech Act Approach**

In this section I review the discussion so far in light of the above account of moral speech acts and discuss the advantages of the speech act model with regard to understanding the connection between moral judgements and motivation. I conclude by raising some potential problems for the speech act approach.

I said in chapter one that the problem of moral motivation results from an apparent tension between the fact-stating and the motivating aspects of moral judgement against the background of a particular theory of motivation, known as the desire/belief theory. The non-cognitivist anti-cognitivist argument is one way of representing this tension. It asserts that the truth of moral judgement internalism entails the falsity of moral cognitivism. I claimed that the anti-cognitivist argument is invalid, and suggested an alternative disjunctive conclusion that allows for the truth of both moral judgement internalism and moral cognitivism. However, I stressed that we should take seriously the non-cognitivist claim that moral judgements in some sense go beyond bare description. I suggested we think of moral judgement internalism as an expression of our interested or participatory relation to moral deliberation.

The task then became one of finding out how best to construe the relation between the fact-stating and the motivating aspects of moral judgement. Content internalism about motivation is one way theorists have attempted to understand this relation. It aims to secure an internal connection between moral judgements and motivation by including the motivational requirement at the level of the truth-conditional content of moral judgements. I rejected content internalism as an interpretation of moral judgement internalism because it threatens the categoricity of moral requirements and the status of internalism as a commonplace of ordinary moral thought. Moreover, it cannot secure the necessary connection between moral judgements and motivation



asserted by moral judgement internalism. Instead, I suggested interpreting moral judgement internalism as a form of *act internalism*, a position that derives from the positive part of the non-cognitivist analysis of moral judgements.<sup>21</sup> Act internalism says that there is a necessary connection between the act of making a moral judgement and motivation. This is perfectly consistent with there being no connection between the truth of moral requirements and motivation. Act internalism is therefore compatible with the widely accepted doctrine that moral requirements apply to agents regardless of their contingent desires.

In the second half of chapter one, I described how Nagel's version of internalism avoids the counterintuitive consequences of content internalism (about motivation) by interpreting internalism as a doctrine about the relation between moral requirements and justifying reasons. I argued that Nagel's conception of the relation between moral judgements and motivation is not a version of moral judgement internalism, as he denies that moral judgements entail motivation. Rather, Nagel advances a cognitivist theory of moral motivation according to which moral judgements normally, but not necessarily, motivate agents to do what they judge to be morally required. Nagel rejects the desire/belief model and explains the motivating force of moral judgements in terms of their justifying content. I argued that if moral considerations play a genuine role in deliberation about what to do, then it must be possible that agents are sometimes motivated by the justifying content of their moral judgements.

Moral judgement externalists argue that in order to secure the fact-stating and cognitive nature of moral judgements, the internalist requirement must be rejected. One reason externalists view internalism as a threat to cognitivism and, by extension, to realism in ethics is that they accept the validity of the non-cognitivist anti-cognitivist argument. However, externalism is much less compelling once the anti-cognitivist argument is removed as an obstacle to the union of cognitivism and internalism. Moreover, externalist explanations of the practical or action-guiding

---

<sup>21</sup> The negative or critical part of the non-cognitivist argument is that the motivating function of moral judgements precludes their functioning to state facts and express beliefs.

function of moral judgements are unconvincing, given their claim that the motivating force and justifying authority of such judgements are only contingently related to moral considerations. I argued that the desire/belief thesis is ultimately a weakness rather than a strength of externalism. A revised version of Smith's fetishism objection to externalism illustrates this point. It claims that the externalist agent is only ever motivated to do what is right in so far as doing so satisfies an independent and contingent desire to do the right thing, and argues that such an agent does conflict with our notion of the virtuous person. This is because our conception of a virtuous agent is of someone reliably motivated by the consideration *that something is right* rather than by a desire to do what is right. For these reasons, I argued we should prefer some alternative account of the relation between moral judgements and motivation to that given by externalism.

Act internalism allows us to avoid locating the connection between moral judgements and motivation at the level of truth-conditional judgement content or at the level of an independent desire to do the right thing. The speech act approach is one way of interpreting act internalism. Speech acts with more than one direction of fit and more than one point or purpose to their performance are ubiquitous in our linguistic practice. Construing moral judgements as a type of linguistic act enables us to account for their fact-stating and motivating aspects. Searle's theory of speech acts distinguishes between the illocutionary force and the propositional content of an utterance. I argued in favour of interpreting the internalist requirement on moral judgements as a claim about their illocutionary force rather than their propositional content. Speech act theory also provides a structured way of understanding the kinds of conditions that govern the performance of different types of linguistic acts. I suggested that moral judgements can usefully be thought of as a type of compound speech act: a speech act that involves the direct performance of more than one kind of illocutionary act, unified by a particular goal. Specifically, I said we should view moral speech acts as evaluations rather than as bare assertions. On this model, moral evaluations involve the direct performance of assertive and commissive illocutionary acts, allowing them to accommodate both the fact-stating and the motivating aspects of moral judgement.

One of the principal advantages of the speech act approach is that it accounts for the sense in which moral judgements go beyond mere description. Moral judgements involve the performance of a commissive as well as an assertive illocutionary act. In performing a commissive illocutionary act, an agent commits herself to acting in a certain way in the future. And in performing a commissive act as part of a moral evaluation, an agent commits herself to being guided by the evaluative content of the proposition, the truth of which she asserts. The essential condition for moral evaluations reflects the sense in which agents making genuine moral judgements are interested participants in the practice of moral deliberation. The essential point of a moral evaluation is to be guided by the content of a proposition that represents those features of a situation that favour acting in one way rather than another. So in making a moral judgement, a speaker expresses her participation in a particular kind of evaluative practice with the specific goal of ranking options to determine how to act.

The speech act approach can also account for the way in which motivation reliably tracks moral judgements. If you convince me to vote for the social democrats rather than the libertarians, I come to judge that it is right to vote for the social democrats. In so judging, I commit myself to acting in accordance with the content of the uttered proposition: that it is right to vote for the social democrats. If my judgement is sincere, I will intend to be guided by this content and so be motivated to act in accordance with it. There is thus a straightforward explanation of the reliability of moral motivation in terms of the essential point of moral judgements.

As it stands, I think the speech act model provides a good account of many of the intuitions underpinning our notion of moral judgement. However, a number difficult issues remain. For example, the model must account for the kinds of failures of motivation resulting from depression and despair mentioned by Stocker. Stocker's argument appears to undermine the strong internalist claim that moral judgements entail motivation. Now, the speech act account is not committed to *this* strong internalist claim. This is because it allows for the possibility that insincere moral judgements – those judgements we don't intend to be guided by – still count as

genuine moral judgements. However, according to the model given above, sincere moral judgements necessarily involve an intention to act in accordance with the content of the uttered evaluation. A problem for this approach is that cases involving severe depression and the like are just those cases where we would expect agents to *fail to intend* to do the actions favoured by their evaluative judgements. It seems reasonable to assume that if someone intends to perform a particular act, then she believes that she will perform that act. Presumably, depressed agents are often all too aware of their lack of motivation, and are unlikely to believe that they will perform the required act. Consequently, in such cases they will have no intention to perform the act in question, even though they may believe they are morally required to perform it. The speech act account needs to explain in what sense an agent can be sincere in her commitment to a moral requirement even when she has no intention of doing what she takes to be right.

As a first step towards dealing with the difficulty of failures of motivation, we might note that to intend to act in accordance with or to be guided by a consideration is not the same as holding oneself to account with regard to a consideration. That is, it seems possible to hold oneself to account with regard to a consideration without intending to act on that consideration. This is the style of response I develop in chapter four, where I use the notion of *normative expectation* to explain the sense in which agents hold themselves to account when they make moral judgements. As a consequence, I make a number of changes to the list of necessary conditions for the performance of moral evaluations. These changes contribute to a more satisfactory account of the nature of the commitment involved in making a moral judgement.

The other main issue the speech act approach needs to address is the relation of moral language to moral thought. It might be argued that while our linguistic conventions entail that the sincere performance of a moral evaluation requires the presence of a motivating attitude, this is just a fact about how we happen to use words; it does not address the deeper question about the connection between psychological assent to moral propositions and motivation. For example, a society may be possible in which it is not linguistically inappropriate to use moral terms and

expressions in the absence of motivating attitudes. I address this issue in chapter five. My response emphasises the level at which the kinds of linguistic rules outlined in speech act theory operate on the natures of agents, like us, who are characterised by their capacity for language and for normative deliberation. To a large extent humans are normative animals, and the capacity to engage in normative deliberation is facilitated in us and partly structured by our ability to use language. Consequently, many of the rules and patterns of behaviour outlined in speech act theory are not such that we can simply choose to ignore them without radically altering the kinds of creatures we are and the kinds of practices in which we engage.

## **10. Conclusion**

The speech act approach should be assessed in terms of its success in accounting for the most important features of moral judgement. Act internalism makes the accommodation of the fact-stating and the motivating aspects of moral judgement *prima facie* plausible. The speech act model suggests a way of interpreting the notion of a moral judgement act as a type of speech performance governed by linguistic rules. It provides a structured way of understanding the relation between such acts and the content of moral judgements in terms of the illocutionary force of a speech act and its propositional content. The internalist requirement is accommodated on this model at the level of illocutionary force rather than at the level of propositional content. Consequently, unlike content internalism, this form of internalism does not threaten the categoricity of moral requirements. The speech act approach provides a detailed account of how moral judgements can be both fact-stating and motivating, and in doing so conforms to deeply held intuitions about the nature of such judgements.

---

## Normative Expectation and the Participant Point of View

### 1. Introduction

In this chapter, I attempt to refine my account of the sincerity condition for the commissive portion of moral judgements given in chapter three. A problem with the intention-based model proposed in that chapter is that it seems unable to account for failures of moral motivation resulting from depression, listlessness, and other kinds of mood disorders. An adequate version of moral judgement internalism must explain in what sense an agent can be sincere in her commitment to a requirement she judges to obtain, even when she has no intention of doing what she takes to be required of her. I argue that an agent is sincere in her commitment to a moral requirement if, in addition to believing it obtains, she normatively expects herself and others to act in accordance with the requirement. This involves her being susceptible to a range of sanctioning attitudes and emotions that make up, what Peter Strawson terms, ‘the participant point of view’ (1962). It is in relation to such attitudes that an agent holds herself and others to account with regard to moral requirements. And it is in this sense, I argue, that making a moral judgement goes beyond merely believing something to be the case.

### 2. Failures of Motivation and Sincerity Conditions

The issue I raise in this section is whether the linguistic approach developed in the previous chapter results in a version of internalism that is too strong, and that, as such, it inherits a number of weaknesses of standard versions of internalism. The concern is that it will prove unable to account for failures of motivation resulting from the kinds of depression and mood disorders that can deprive an agent of the motivation to do what her evaluative judgements recommend.

As we have seen, Stocker argues persuasively that changes in the mood, interest, and energy of agents as a result of depression and the like can often mean that agents fail to be motivated by what they genuinely judge to be good (1979). Stocker’s argument



seems to undermine the strong internalist claim that failures of motivation necessarily invalidate moral judgements. As already mentioned, the speech act approach is not committed to the strongest version of the internalist thesis: that all failures of motivation, even those due to insincerity, render unsuccessful the performance of a moral judgement. This is because on the speech act account insincere moral judgements still count as genuine moral judgements. Recall that on that account a speaker is sincere in performing a moral speech act when she *intends* to act in accordance with the content of the moral proposition expressed by her utterance.<sup>1</sup> If the agent has no intention of acting in accordance with the uttered moral proposition, then her judgement is insincere. Nonetheless, she will still have succeeded in making a moral judgement and will still have expressed her commitment to being guided by the uttered moral proposition. The speech act approach can thus accommodate the possibility of genuine moral judgements that fail to motivate.

This move, however, does not help the speech act approach in relation to Stocker's argument, because his argument also undermines the claim that *sincere* moral judgements necessarily motivate. Stocker's point is not that the depressed agent makes a genuine though insincere moral judgement, but is the stronger claim that the depressed person is perfectly sincere in her moral judgements despite her lack of motivation. The speech act model says that sincere moral judgements necessarily involve a motivating intention to act in accordance with the content of the uttered evaluation. The difficulty for this model is that cases in which agents are suffering from severe depression and the like are just those cases where we would expect agents to fail to intend to do the actions favoured by their evaluative judgements. It seems reasonable to assume that if someone intends to perform a particular act, then she believes she will perform that act.<sup>2</sup> Presumably, depressed agents are often all too aware of their lack of motivation, and are unlikely to believe that they will perform the required act. Consequently, in such cases they will have no intention to

---

<sup>1</sup> As we have seen, an agent must also believe the moral proposition expressed by her utterance in order to make a sincere moral judgement. In making the following comments I assume this condition is satisfied.

<sup>2</sup> Paprzycka (1999) notes that this position is defended by Grice (1971) and Velleman (1989) but called into question by Davidson (1980).

perform the act in question, even though they may believe they are morally required to perform it. And this looks like a potential vindication of externalism. In order to provide an adequate defence of moral judgement internalism, the speech act approach must explain in what sense an agent can be sincere in her commitment to a judged requirement even when she has no intention of doing what she takes to be right, *and in what sense this commitment goes beyond mere belief*. What is needed is an account of the sincerity condition for the commissive portion of a moral speech act capable of accommodating the kinds of failures of motivation outlined by Stocker.

There are a number of other reasons why intention is a problematic state to fill the role of sincerity condition for the commissive portion of moral judgements. Take the simple example of weakness of will. Suppose I judge that I am morally required to help at the local library fundraiser, but on the day it occurs find I am not sufficiently motivated to leave the warmth and comfort of my home. I feel guilty about not helping but cannot pull myself away from the pleasures of drinking coffee and reading the papers. Now, in this case the motivation I undoubtedly have to help at the fundraiser is overwhelmed by my desire to stay warm and drink coffee. I have no intention of helping at the fundraiser and every intention of staying home and enjoying myself. Nonetheless, it still seems that I can be perfectly sincere in my judgement that I am morally required to help at the fundraiser, and still, in some sense, sincere in my commitment to the requirement expressed by my judgement.

Intention as a sincerity condition also poses a problem in relation to moral judgements that call for agents other than the speaker to act in certain ways, and in ways that are not directly relevant to, or possible for, the agent making the judgement.<sup>3</sup> Suppose that John believes that all women should breast-feed their babies rather than bottle-feed them. Now, while John is committed to women breast-feeding their babies, it is obviously not something he can intend to do himself.

---

<sup>3</sup> Moral judgements involving supererogatory acts also seem to pose a problem for the notion of intention as a sincerity condition for moral judgements. For example, I may sincerely judge a soldier's behaviour as morally admirable because courageous, without thereby intending to perform such actions myself.

Despite this, it may seem that the version of intention as a sincerity condition already given can accommodate this kind of problem. Recall, I said that an agent is sincere in his commitment to a moral judgement insofar as he intends to be guided by the content of the judgement. Perhaps John intends to be *guided* by his belief that women should breast-feed their babies, even though he cannot engage in breast-feeding babies himself. For example, he might actively promote breast-feeding and encourage those close to him, his wife and female friends, to breast-feed rather than bottle-feed their babies. John is also very likely to express his disapproval of women who opt to bottle-feed where breast-feeding is an option, and to commend those who follow his advice and breast-feed their babies. This suggests that whether an agent is sincerely committed to a principle of action or to a behavioural standard is linked, in some sense, to how she would respond and the emotions she would feel in situations where the principle or standard is violated or upheld. It is not clear, however, that the notion of an intention, even an intention to be guided in some way, is adequate to capture these more complex types of responses.

As a way forward, we might note that to intend to act in accordance with a consideration is distinct from holding oneself to account with regard to a consideration. I may intend to act in accordance with the consideration that it is right to give money to Oxfam, but only do so in order to win favour with my colleagues. In such a case it seems incorrect to describe my intention as evidence of a sincere commitment to the principle that it is right to give money to Oxfam. By contrast, it seems possible to be committed to a consideration in the sense of holding oneself to account with regard to the consideration, without intending to act on it. For example, suppose that John becomes deeply depressed because he cannot convince the women in his life to breast-feed rather than bottle-feed their babies. This causes him to stop promoting breast-feeding and to stop lecturing his friends and relatives about its superiority to bottle-feeding. Nonetheless, John remains committed to the view that women have an obligation to breast-feed rather than bottle-feed their babies. He still holds himself to this view even though he now lacks any intention to act in accordance with it. Let's say that an agent holds herself to account with regard to a consideration or requirement if, among other things, she understands and accepts that

the consideration or requirement applies to her and that, as such, she can be justifiably criticised for failing to comply with it. In the case of a sincere moral speech act, then, the speaker understands and accepts that the moral requirement applies to her and that, as such, she can be justifiably criticised if she fails to satisfy the requirement. More succinctly, we might say that in making a sincere moral judgement an agent *normatively expects* herself to act in accordance with the requirement expressed by the judgement.

How should we characterise the notion of normative expectation? A useful contrast is sometimes made between normative expectations and predictive expectations (Paprzycka 1999: 631). Predictive expectations are said to be belief-like. For example, if I predictively expect Sarah to be home before 5pm, then I believe that she will be home before that time. Normative expectations, by contrast, seem to be demand-like. If I normatively expect Sarah to be home before 5pm, then I place a demand on her to be home before 5pm: I hold her to account with regard to the expectation. According to an account offered by Katarzyna Paprzycka, this involves my being disposed to sanction negatively her failure to be home before this time and to sanction positively her success in being home on time (1999: 632). A crucial difference, then, between predictive and normative expectations is that the latter but not the former must be had towards responsible subjects of action. In general, if I can believe that *p* will happen in the future, then I can predictively expect that *p*. If I can believe that it will rain tomorrow, then I can predictively expect that it will rain tomorrow. By contrast, I cannot normatively expect that it will rain tomorrow, but I can normatively expect that you pick me up from work if it rains tomorrow.

In addition to having normative expectations of others it seems that we can have normative expectations of ourselves. If I normatively expect myself to get to the meeting on time, I place a demand on myself to get to the meeting and will be disposed to sanction myself if I fail to meet this demand. One suggestion is that the kind of expectation involved in moral judgement involves this kind of self-regarding normative expectation. Using this notion, we might explain the possibility of an agent who sincerely judges that an act is morally right, yet has no intention of acting

in accordance with his judgement, and believes he will not so act. The crucial feature to note about normative expectation is that it involves a demand that does not have to be withdrawn if it fails to be satisfied. Contrast this with predictive expectation. If I predictively expect you to be home before 5pm and you in fact get home at 7pm, then my predictive expectation is at fault and must be 'withdrawn.' If I normatively expect you to be home before 5pm and you get home after this time, my normative expectation will continue to apply to your behaviour and need not be withdrawn: it is your behaviour rather than the expectation that is at fault.<sup>4</sup>

We can apply this account to the case of the depressed agent who makes a moral judgement (performs a moral speech act) but who fails to be moved and fails to intend to be moved to act in accordance with it. John announces to his wife that he has a moral obligation to vote in the national election. Unfortunately, he suffers from depression and so is not motivated to go to the polling booth and fulfil his obligation. He doesn't intend to go and doesn't believe that he will go and cast his vote. According to the account offered above, John's moral judgement is sincere so long as he believes that it is right for him to vote in the election and normatively expects himself to act in accordance with the proposition that it is right for him to vote. The essential condition for moral judgements means that issuing his moral utterance commits John to acting in accordance with the uttered proposition; issuing it *sincerely* necessitates that he holds himself to account with regard to the content of the proposition.

---

<sup>4</sup> This will remind us of 'direction of fit' distinctions typically used to distinguish belief-like states from desire-like states. Paprzycka notes that predictive expectations can be characterised as having a belief-like 'mind-to world' direction of fit whereas normative expectations can be characterised as having a more desire-like 'world-to-mind' direction of fit (1999: 632). Below I question the claim that normative expectations are wholly desire-like in their direction of fit.

### 3. Normative Expectation and Universalizability

Unfortunately, the notion of self-regarding normative expectation does not account for the way in which agents hold others as well as themselves to moral demands. Not all cases of holding someone to account or normatively expecting something of them are moral in character. And purely self-regarding normative expectations seem to be of this non-moral type. Suppose I commit myself to losing a certain amount of weight by next month and so hold myself to the expectation that I lose the weight. This expectation or demand does not seem to generalise to others in the way it would if it were a moral consideration. An important feature of specifically moral expectations is their applicability to moral agents in general: moral expectations inherit the universalizability of moral requirements or demands. This means that if it is wrong for a particular agent to  $\Phi$  in circumstances C, then it is wrong for any relevantly similar agent to  $\Phi$  in C (see Mackie 1977: chap. 4; Hare 1981: chap. 6).

The generalisable nature of moral requirements means that moral judgements commit those making them to holding the same view about relevantly similar actions and agents in relevantly similar circumstances. If I judge that I am morally required to return the money to the bank, this commits me to the view that anyone relevantly similar to me in similar circumstances would be morally required to return the money to the bank. The universalizable nature of moral requirements means that my normative expectation cannot simply be limited to my own conduct. This would have the unfortunate consequence that I only ever held myself to account with regard to moral requirements. And I think that if an agent in this way failed to grasp the relevance of universalizability to moral assessment, we would be unlikely to view her as capable of making genuine moral judgements. A moral judge must be 'ready to apply [moral requirements] equally to himself and to others, and to go on applying in interpersonal situations when the roles are reversed...' (Mackie 1977: 85).

We might put this point another way. A moral judge does not merely hold herself to account in sincerely issuing a moral judgement, she holds to account all those who she deems are appropriate objects of moral appraisal. If I judge that I am morally required to vote, then I normatively expect all morally responsible agents relevantly



similar to me in circumstances relevantly similar to mine to vote. That is, I hold myself to account for the same reason I hold other moral agents to account: the universal applicability of moral requirements to those who are appropriate objects of moral appraisal. The formulation of the sincerity condition for the commissive portion of moral judgements needs to accommodate the universalizability of moral requirements. It can do this by extending the scope of the normative expectation involved in moral judgement to include all who are the appropriate objects of moral appraisal. Thus we get the following formulation:

If S sincerely utters the judgement that  $\Phi$  is morally right, then S normatively expects all those who are appropriate objects of moral appraisal to act in accordance with the proposition ‘ $\Phi$  is morally right.’

On the account of normative expectation given above, the new formulation means that an agent making a moral judgement is disposed to sanction negatively a failure to comply with the judged requirement and to sanction positively conformity with the requirement by all agents to whom the judgement correctly applies.

An important difference between the agent who makes the judgement and the other agents to whom the judgement applies is that the judging agent has responsibility for her own actions in relation to the requirement, whereas, in general, she is not responsible for the actions of others. Thus the forms of sanctioning responses and emotions the judging agent will apply to herself will often be different to those she applies to others. For example, she will tend to feel guilt or shame when she believes she has violated a requirement and moral disapproval or indignation in response to transgressions she believes others have made. However, the general stance of holding someone to account or normative expectation is essentially the same whether the agent directs the expectation towards herself or to others. This neatly reflects the irrelevance of numerical difference characteristic of the process of universalization: the mere numerical fact that I am who I am and that you are who you are is irrelevant to the applicability of moral demands to us (cf. Mackie 1977: 83-90).

#### 4. Unwarranted Sanctioning Emotions and Responses

The proposed account of normative expectation identifies the stance involved with a propensity to respond in certain ways to the violation of or conformity with standards of conduct. As described, the stance appears to involve the expression of sanctioning emotions or the response of sanctioning behaviour rather than a judgement of any kind. This would seem to make the stance conducive to the project of accommodating the motivating aspect of moral judgement. However, I argue below that this purely non-judgemental formulation of normative expectation cannot be correct. This is because it allows agents to hold others, and themselves, to expectations in situations they do not believe warrant their having such expectations: situations they do not believe warrant the response of sanctioning behaviour or emotions.

According to Paprzycka's account, a person B normatively expects of another person S that p, when B is disposed to sanction negatively S's failure to bring about p and to sanction positively S's success in bringing about p (1999: 632).<sup>5, 6</sup> She argues that normative expectations have a world-to-mind direction of fit like desires – the world (the person held to account) must conform to the expectation – rather than a belief-like direction of fit. The problem with this account of normative expectation is that it fails to exclude situations in which an agent does not accept as justified the sanctioning response she makes or the sanctioning emotion she feels. For example, John makes disapproving comments whenever Sarah does not come top of the class. He says things such as, 'Well, you did spend a lot of time on the phone,' and, 'I expect you will try harder next time.' John also feels a certain sense of disappointment on the rare occasions that Sarah fails to top the class. When Sarah

---

<sup>5</sup>For ease of expression, I won't always include reference to positive sanctions in my discussion of normative expectation. This merely reflects the fact that it is often when moral requirements are breached rather than when they are adhered to that sanctioning emotions are most vividly expressed. This emphasis is not intended to imply that positive sanctions only have a secondary role in holding agents to account (cf. note 10 below).

<sup>6</sup>Paprzycka continues: 'The notion of sanctions is broad enough to include the reactive emotions [see below for an account of such emotions] as well as sanctions of a lesser moral magnitude like feeling dissatisfied or disappointed by oneself or by another, criticising oneself or others, and, on the side of positive sanctions, feelings of satisfaction and accomplishment' (1999: 632). I follow her in stressing a broad construal of the positive and negative sanctioning emotions and responses involved in holding agents to expectations.

eventually confronts John and explains that his comments make her feel bad, John is quite shocked. He has been making the comments unreflectively and certainly does not believe that Sarah's approach to her work warrants disapproval or sanction. That is, John does not believe that there is a requirement that Sarah always come top of the class, even though he often behaves as if he endorsed such a requirement. On Paprzycka's account of normative expectation, John normatively expects Sarah always to come top of her class if he is disposed to sanction her failure to do so. However, given that John does not actually believe such an expectation justified, it seems odd to say that he really holds her to account with regard to the requirement that she always top the class.

Consider another example.<sup>7</sup> Suppose having been raised in a traditional Catholic home, I have a tendency to feel disapproval towards those who engage in sexual relations outside the institution of marriage. However, let's say that I have come in my adult life to reject the doctrine that it is morally wrong to have sexual relations outside marriage. Consequently, whenever I experience the disapproval in question, I judge it to be completely unwarranted and simply a product of the emotional residue of my childhood experience of Catholicism. This includes those times when I feel guilty for engaging in such behaviour myself. The problem with the account of normative expectation given above is that it does not seem to be able to deal with examples of this kind. On that account, the fact that I am disposed to sanction negatively engaging in sex outside marriage means that I normatively expect myself and all relevant agents to abstain from sex outside marriage. But this means that I hold myself and others to a requirement that I do not in any way endorse.

These examples suggest that merely having a disposition to sanction a certain form of behaviour is not sufficient to hold someone to account with regard to a requirement prohibiting such behaviour. It would seem that for an agent to normatively expect someone to act in a certain way, the agent must believe that the expectation or requirement in question is warranted, in addition to being disposed to sanction a failure to conform to the expectation. Now, it may be that this belief-like

---

<sup>7</sup> Wallace uses a similar example to explain the notion of unwarranted guilt (1994: 43).

or judgemental feature of normative expectation follows directly from the nature of the kinds of attitudes or emotions typically expressed in sanctioning practices. Following Strawson, the kinds of attitudes involved in blaming and sanctioning practices – resentment, guilt, moral disapproval, moral indignation, and the like – have been termed ‘reactive attitudes’ (1962)<sup>8</sup>

These attitudes are ‘reactive’ in that ‘they are reactions to the moral qualities exemplified by an individual’s attitudes and conduct’ (Watson 1987: 122). The reactive attitudes therefore have ‘internal criteria’ according to which they are or are not appropriately held by an agent (Watson 1987: 122). For example, if one feels guilty it seems there must be some perceived fault in one’s behaviour towards which the guilt is directed. Similarly, if one feels moral indignation towards someone, there must be something, some wrong, they have done that one is indignant about.

Wallace interprets this feature of the reactive attitudes as their having a distinctive kind of propositional content (1994: 50)<sup>9</sup> Specifically, he argues that any state of negative reactive emotion, such as guilt or resentment, entails that the agent who has the state believes that some expectation or requirement has been breached.<sup>10</sup>

If being in a state of reactive emotion entails a belief that a requirement has been violated, then it might seem we can rest content with our initial formulation of normative expectation. Assume that we interpret sanctioning responses in terms of the expression of reactive emotions. To normatively expect Sarah to come top of the class is to be disposed, say, to feel resentment towards her if she fails to come top of the class; and to feel resentment towards her entails that one believes she has violated

---

<sup>8</sup> I will use the expressions ‘reactive attitude’ and ‘reactive emotion’ interchangeably.

<sup>9</sup> As we will see below, Wallace’s account differs from Watson’s in holding that although the propositional content of reactive emotions involves the notion of a breached requirement or expectation, the requirement in question need not be specifically moral in character nor one that the agent experiencing the emotion necessarily endorses (see Wallace 1994: 20).

<sup>10</sup> Wallace does not think that holding agents morally responsible for worthy acts is essentially linked to positive sanctioning attitudes in the way that holding agents responsible for unworthy acts is essentially linked to negative sanctioning attitudes. He attempts to give an account of what it is to hold someone responsible for worthy acts in terms of what he sees as the more fundamental connection between holding someone responsible in general and the negative sanctioning attitudes (Wallace 1994: 71). As I say in note 5 above, I don’t think that the prominence of the negative sanctioning attitudes in discussions of normative expectation necessarily signals a theoretical asymmetry of the kind Wallace proposes.

some requirement – presumably, the requirement that she tops the class. This interpretation of the reactive emotions would simply exclude the possibility that one can feel resentment or guilt in the absence of a belief about a violated requirement. So if an agent *feels* guilty about engaging in sexual relations outside marriage, but does not believe that he has violated any requirement prohibiting such behaviour, then he is not experiencing a genuine state of guilt. Perhaps he merely experiences feelings that resemble guilt in some respects.<sup>11</sup> So, on this model, a disposition to respond with the reactive emotions is a disposition to respond to believed violations of requirements with certain emotions: the genuine experience of reactive emotions simply entails such beliefs.

How plausible is this type of judgemental view of the reactive emotions? Wallace himself concedes that we do seem to experience genuine but irrational guilt and resentment (1994: 40, 46-47). That is, we sometimes feel real guilt and real resentment we do not believe to be warranted. Wallace aims to accommodate this phenomenon while maintaining the view that the reactive emotions entail beliefs of some kind (1994: 40-50). This is necessary, he argues, in order to capture the kind of propositional content that is distinctive of such emotions. His approach is to distinguish between the belief that some requirement has been violated, the belief entailed by a reactive emotion, and the belief that the violation of a requirement *warrants* one of the reactive emotions. Wallace aims to provide a middle-way between accounts of the reactive emotions in terms of exclusively moral beliefs and accounts that claim that the reactive emotions can be understood independently of cognitive states. The first of these ‘overmoralises’ the account of the reactive emotions by insisting that the explanation of any state of reactive emotion must

---

<sup>11</sup> As Wallace notes, Rawls defends a view along these lines (see Wallace on Rawls 1994: 47; Rawls 1972: 481-482; and note 12 below.)



invoke a moral concept (Wallace 1994: 19, 46)<sup>12</sup> This view is difficult to sustain, suggests Wallace, as agents often experience such emotions without believing they have breached a moral requirement or that they are morally accountable. The approach which eschews the role of belief, by contrast, fails to account for the distinctive propositional content of reactive emotions that connects them essentially with the perceived violation of demands (Wallace 1994: 19, 49).<sup>13</sup> Wallace's generic stance of 'holding someone to an expectation' is meant to provide the middle way between these two extremes. It is what he terms a 'quasi-evaluative' stance rather than a full blown stance of moral endorsement:

To be in a state of reactive emotion, one must believe that a person has violated some expectation that one holds the person to: and in terms of this belief, we can give an account of how the reactive emotions have the kind of propositional content that distinguishes them from other emotional states. But it need not be the case that the expectation that gives the content of a reactive emotion is a moral one, or even that it is an expectation one sincerely endorses. (Wallace 1994: 20).

Wallace accepts that the kind of belief involved in his stance of holding someone to an expectation only counts as a belief in a degenerate sense: '[A] state in which one entertains a proposition one does not fully accept – to which one assigns an extremely low probability, for instance – may be allowed to count as a belief' (1994: 24). The special nature of this quasi-evaluative stance is intended to explain how reactive emotions can be essentially explicable in terms of beliefs about the violation of requirements, thus accounting for the distinctive propositional content of such emotions, while allowing that agents can experience a genuine reactive emotion without believing they have committed any moral infraction or that the emotion is a

---

<sup>12</sup> This is the position Wallace attributes to Rawls and which informs Rawls' approach to unwarranted reactive attitudes (see Rawls 1972: 481-482). It is worth distinguishing between Rawls' claim that reactive attitudes such as guilt and indignation must be explained in relation to beliefs about specifically moral infractions and the claim that the reactive attitudes must be explained in relation to normative beliefs more generally. On the latter view, in order for an agent to feel genuine, say, resentment, she must believe that someone has behaved in a way that *warrants* the emotion of resentment; she need not believe that the person has done anything morally wrong. While the charge of overmoralizing the reactive attitudes is avoided by this latter model, it still leaves the problem of explaining why reactive emotions that an agent experiences but doesn't believe warranted do not count as genuine instances of such emotions. Wallace rejects explaining reactive emotions in terms of specifically moral beliefs and rejects explaining them in terms of normative beliefs more generally. As I outline in the text, he argues that any given reactive emotion must be explained in terms of the 'quasi-evaluative' belief that some expectation or requirement has been violated.

<sup>13</sup> Wallace attributes such a view to Gibbard (see Wallace 1994:48-50; and Gibbard 1990: 135-150).



warranted response to the situation in question. On Wallace's model, then, it is possible for unwarranted reactive emotions to be examples of holding someone to an expectation because holding someone to an expectation does not entail that one endorses the expectation or believes that its violation would warrant a negative sanction.

There are a number of problems with Wallace's quasi-evaluative version of holding someone to an expectation. For instance, it would seem to be vulnerable to the same kind of objection made against purely non-judgemental accounts of normative expectation. According to Wallace, one can hold someone to an expectation without believing the expectation warranted and without believing that a failure to conform to it would warrant a negative sanction. But it seems inaccurate to our ordinary notion of holding someone to account that we would do so while believing that there were no grounds on which to base our stance. The apostate may feel guilt for engaging in sexual relations outside marriage, but he does not hold himself to a requirement to abstain from sex outside marriage precisely because he does not believe such a requirement warranted.

The quasi-evaluative version of holding someone to an expectation also seems to get the propositional content of the reactive emotions wrong. Wallace argues that the reactive emotions are distinguished by their 'presumptive' connection to expectations (1994: 20-25, 31, 39). Specifically, they are distinguished by the propositional content that characterises them as a class. This is the content of the 'degenerate' belief that some expectation has been violated (Wallace: 1994: 24). But this content is ill-suited to be the defining feature of the reactive emotions, as it points to an item that is only nominally an expectation or requirement. And purely nominal requirements do not call for the response of a reactive emotion. For example, the non-Muslim who believes that a woman has violated an Islamic 'requirement' to cover her hair in public does not believe that there is an actual requirement (a normative requirement) that women cover their hair in public. He believes that there is a standard of conduct that a certain group treats as a normative requirement, and that this standard has been violated. But the content of this belief, by its very nature,

does not justify or make appropriate the response of a reactive emotion. By contrast, an agent's belief that the woman's failure to cover her hair in public *warrants* moral indignation allows us to see the indignation, from the agent's perspective, as an appropriate response to the violation of a normative requirement.

It would seem that reactive emotions only ultimately make sense in terms of central cases in which they are *warranted* responses to the violation of or conformity with expectations. Agents understand what indignation is because they understand it as an emotion that is only appropriately held in circumstances where a justified requirement or standard of behaviour is breached. This suggests that the propositional content characteristic of the reactive attitudes is essentially evaluative rather than quasi-evaluative. How then do we account for the fact of unwarranted reactive emotions? I suspect that the propositional aspect of unwarranted reactive emotions can be accounted for without invoking the controversial notion of a 'degenerate' belief. As Wallace himself notes, susceptibility to a particular reactive emotion, like guilt, in a certain context will be accompanied by a range of evaluative thoughts one does not necessarily endorse (1994: 46). If I feel guilty about her leaving early, it is perhaps because I wonder whether I should have been more welcoming to her, or that I said something to upset her. Now, suppose that after deliberation I conclude that the guilt I feel is not warranted, because I was sufficiently hospitable to her and said nothing to upset her. We can still explain my feeling guilty by noting the kinds of demands and considerations that I *thought about* on her leaving. Sometimes, taking seriously or thinking about the prospect that one has done wrong may be enough to trigger feelings of guilt, even in the absence of a belief that one has done something wrong or violated a requirement. In such cases, the presence of the guilt is still made sense of in relation to the notion of particular requirements whose violation, if they obtained, would make the guilt *appropriate*. In this way, the distinctive propositional content of reactive emotions is preserved as an essential part of their explanation without the need to posit an actual belief about the violation of a requirement.

Wallace's stance of 'holding someone to an expectation' is not robust enough to play the role I need the notion of normative expectation to play. Recall that on my account normative expectation is meant to be a sincerity condition for the commissive portion of the speech act of moral evaluation. That is, it determines whether an agent is *sincere* in her commitment to being guided by the content of her moral judgement. If it were possible for an agent to believe a requirement unjustified and yet normatively expect someone else to conform to the requirement, then the fact that the agent has this normative expectation could hardly be an expression of her sincere commitment to the requirement in question. The robust sense of normative expectation required by my approach means that our initial formulation of it remains inadequate because it makes unwarranted sanctioning responses and emotions sufficient for holding someone to an expectation. Given these considerations, it seems that a satisfactory conception of normative expectation must make reference to a requirement that the agent who holds the expectation accepts or believes is warranted and believes warrants appropriate sanctions if violated.

##### **5. Holding Someone to an Expectation and the Reactive Emotions**

In the last section I argued against Wallace's claim that any experience of a reactive emotion involves the stance of holding someone to an expectation. Wallace defends this view because he thinks that the notion of holding someone to account must be made sense of in terms of the reactive emotions. This is part of a broader attempt to develop an account of moral responsibility based on Strawson's views in 'Freedom and Resentment' (Wallace 1994: 18).<sup>14</sup> Wallace commits himself to what he takes to be one of Strawson's central themes: that moral responsibility should be understood in terms of susceptibility to the reactive attitudes. Using this theme as a starting point, Wallace proceeds 'by working out an account of the stance of *holding* someone responsible in terms of the reactive emotions' (Wallace 1994:18. My emphasis).

---

<sup>14</sup> See Strawson 1962.

Strawson's central concern in 'Freedom and Resentment' is the issue of moral responsibility with regard to the free-will and determinism debate (1962). Strawson defends a non-consequentialist form of compatibilism about moral responsibility, designed to show that the thesis of determinism has no bearing on the propriety of the practice of holding people morally responsible – a practice characterised, he claims, by an essential connection to the reactive attitudes.<sup>15</sup> My focus here is not primarily what it is to *be* morally responsible but is the nature of our practice of *holding* people responsible or accountable for their behaviour. On Strawson's view, there are standards internal to our practice of holding agents responsible that indicate when it is and is not appropriate to hold someone to account (1962: 50-52). His argument suggests that there is nothing more to our notion of being morally responsible than is captured by the standards involved in holding people morally responsible. But one need not accept this aspect of Strawson's position in order to support the broadly Strawsonian theme that *holding* someone morally accountable is essentially connected to susceptibility to the reactive emotions.

The concept of holding someone responsible or accountable suggests a level of engagement or interest beyond that involved in simply believing or describing something (cf. Wallace 1994: 52). Indeed, the way in which normative expectation goes beyond cognition is, I contend, precisely what is distinctive about such a stance. Consider a world in which to hold someone accountable involves merely noting that the person has done some wrong that warrants a negative sanction, but where the stance involves no tendency to enforce the sanction or to experience sanctioning emotions. When someone from this world says, for example, 'I hold you to account for losing the money,' she is not in the least inclined towards blaming the person for losing the money or engaging in any practical sanctioning measures against the

---

<sup>15</sup> I am not here concerned with the issue of whether determinism is relevant to the justification of the concepts of moral responsibility and freedom and related practices. If it is possible to make genuine moral judgements (in a non-revisionary sense), then it must be possible for agents to be responsible and to be held responsible for their actions. If the kind of incompatibilism which denies the possibility of moral responsibility, and thus the propriety of moral sanctions, is correct, then it is never the case that agents make genuine moral judgements, and so the issue of moral judgement internalism becomes irrelevant. I am assuming, therefore, that agents sometimes makes genuine moral judgements, and that it is possible for agents to be responsible and to be held responsible for their actions.

person, even though she may judge that the person ‘warrants’ such measures. Moreover, she sees nothing unusual or amiss in this dislocation between what is called for and what is done, for holding someone to account is just to make a disinterested observation about the behaviour of an agent. I take it that what this person means by ‘holding someone to account’ is something quite different to what we normally mean by the expression. As Gary Watson notes, to hold someone to an expectation ‘means something in practice...[It] is to be ready to treat them in certain ways’(1987: 120). I agree with Watson that any adequate analysis of our ordinary notion of holding people accountable must capture this practical aspect of the stance.<sup>16</sup>

According to Wallace, the role of the reactive emotions is crucial to account for the ‘form of deep assessment that goes beyond mere description’ characteristic of the stance of holding someone responsible (1994: 82). It is thus vital that his theory captures the nature of the essential connection between the reactive attitudes and the stance of holding someone responsible. Wallace characterises the connection in the following way:

[T]he stance of holding someone to a demand...does not have explanatory priority vis-à-vis the reactive emotions: to be subject to the reactive emotions is to take this stance towards a person, and to adopt this stance is in turn to be subject to the reactive emotions. (Wallace 1994: 24)

I will refer to the claim expressed in this statement as the ‘mutual dependence claim’. It asserts that the notion of holding someone accountable and the notion of susceptibility to the reactive emotions can only be elucidated in terms of each other.

It is clear that Wallace’s quasi-evaluative stance of holding someone to an expectation is developed to accommodate the Strawsonian insight captured by the mutual dependence claim. This special stance is needed to counter the appearance that the mutual dependence claim is contradicted by the fact of unwarranted reactive emotions. I have already argued that it is awkward to describe an agent who feels a

---

<sup>16</sup> I will use concepts of ‘holding to account’, ‘holding to an expectation or requirement’, and ‘normatively expecting someone to conform to a requirement’ interchangeably.

reactive emotion towards someone without believing it warranted as holding that person to an expectation. But this seems to confute the claim that simply being subject to a reactive emotion is to adopt the stance of holding someone to an expectation. By contrast, the quasi-evaluative stance advanced by Wallace allows that an agent can hold someone to an expectation even if the agent does not endorse the expectation or believe that its violation would warrant a negative sanction. On this model, even if I don't believe the guilt I feel about cursing is warranted, if it is genuine guilt, I must believe I have violated a requirement not to curse. This will then count as my holding myself to account with regard to the requirement not to curse. In this way, Wallace aims to accommodate both the mutual dependence claim and the fact of unwarranted reactive emotions.

A problem with Wallace's approach is that it weakens the stance of holding someone to an expectation to such an extent that it no longer resembles the robust stance of holding someone responsible with which he was initially concerned.<sup>17</sup> This forces him to develop a quite different account of what it is to hold someone morally responsible. According to Wallace, an agent only holds someone morally responsible with regard to a requirement if she *accepts* the requirement 'as a basis for practical deliberation and normative criticism and discussion' (1994: 63). Because on this account the notion of 'acceptance' does all the work in terms of elucidating the stance of holding someone responsible, the initial Strawsonian connection endorsed by Wallace between this stance and the reactive attitudes is left inadequately explained. And this means we lose the central insight expressed by the mutual dependence claim.

Wallace interprets the mutual dependence claim narrowly in insisting that any particular experience of a reactive emotion or disposition to experience a reactive

---

<sup>17</sup> This is related to the problem, already mentioned, that Wallace's weakened stance of holding someone to an expectation allows agents to hold others and themselves to expectations they in no way endorse. The point here is that implicit in his development of this attenuated version of the stance is a recognition that it is implausible that an agent would hold someone morally responsible for an act if the agent did not believe there to be any justified moral demand requiring or prohibiting the act. This is why Wallace's conception of holding someone to an expectation comes apart from his conception of holding someone morally responsible.



emotion is an example of holding someone to an expectation. We have already seen that this view faces difficulties in accounting for unwarranted reactive emotions. But it also faces difficulties in accounting for cases in which an agent believes that someone warrants being held to account, but the agent's ability to feel the relevant reactive emotion is inhibited by fatigue, depression, or some other factor. Consider the following example. John is waiting at a bar to be served when a friend of the bartender pushes in and is served before him. John recognises that it would be appropriate to feel resentment towards both the bartender and the bartender's friend, but he is too tired and depressed to be able to work up any feelings of anger or resentment towards them. So he continues to wait patiently to be served. Now, it would seem that despite his failure to feel resentment, John holds the bartender and her friend accountable for their behaviour and believes that such behaviour warrants a negative sanction. But this suggests that, contrary to the mutual dependence claim, it is possible to hold someone to an expectation without being subject to a reactive emotion.

One might expect Wallace to defend the mutual dependence claim by denying that an agent who fails to feel a reactive emotion can be said to hold someone to an expectation. But Wallace argues that it is possible to hold someone to an expectation merely by believing that the person would warrant a negative reactive emotion were she to violate the expectation (1994: 21). To accommodate this view, he proposes a disjunctive account of what it is to hold someone to an expectation. Wallace says that to hold someone to an expectation it is sufficient either that one be susceptible to the reactive attitudes in cases where the person violates the expectation or that one believes that one of the reactive attitudes would be warranted in such instances (Wallace 1994: 21).<sup>18</sup> However, this account would seem to undermine the mutual dependence claim by allowing that one can hold someone to an expectation by way of a belief about the reactive emotions rather than by being subject to such emotions. Wallace recognises this tension, but argues that because the reactive emotions are

---

<sup>18</sup> One might have both a susceptibility to the reactive attitudes and a belief that a reactive attitude is appropriate (Wallace 1994: 23). Note that this latter belief is not the 'degenerate' belief entailed by any experience of a reactive emotion: the belief that some requirement has been violated. It is not clear what the relationship between these two kinds of beliefs – the 'justifying' and the 'degenerate' – is meant to be.

included in the content of such beliefs, someone who has a belief of this kind is sufficiently connected to these emotions (1994: 76,77).

It seems doubtful that Wallace's disjunctive account can capture the sense in which holding someone to an expectation is essentially tied to susceptibility to the reactive emotions. This is because it is precisely an agent's *susceptibility* to such emotions that makes it the case that she goes beyond mere belief or cognition in holding someone to account. What is needed is an account of the essential connection between normative expectation and the reactive emotions that allows for instances of unwarranted reactive emotion and explains cases in which a belief that an emotion is warranted gives the appearance of being sufficient for holding someone to an expectation. To develop such an account, I turn in the next section to Strawson's model of the reactive emotions.

#### **6. Susceptibility to the Reactive Emotions and the Mutual Dependence Claim**

Wallace's view of the reactive emotions emphasises the relation between the stance of holding someone to an expectation and being susceptible to the reactive emotions in particular cases. This approach allows that one can fail to experience a particular reactive emotion yet still succeed in holding someone to account, provided one believes the emotion warranted. In such cases, one is not subject to the relevant reactive emotion or emotions. But this version of the relation between holding someone to an expectation and the reactive emotions fails to capture a crucial strand of Strawson's account of these emotions. It fails to capture the sense in which, for Strawson, to be subject to the reactive emotions is to be *generally* susceptible to a range of emotions and attitudes that characterise a background 'framework' within which moral deliberation and practice take place (1962: 55). On this model, to hold someone to an expectation is to be prone to the reactive or 'participant' attitudes; and to be prone in this way is to be a participant in the practice of assessing and reacting to the attitudes, intentions, and behaviour of agents as they impact upon one and upon others (cf. Strawson 1962: 64). It is to be, what I will term, *normatively engaged*.

In this section I want to outline and defend a broad or general reading of the connection between the reactive emotions and the stance of holding someone to account. This is a reading I take to be implicit in Strawson's account of the reactive emotions and of holding agents morally accountable. As I have said, Wallace's adopts a narrow construal of the mutual dependence claim. He insists that any genuine experience of a reactive emotion involves the stance of holding someone to account. This approach leads to the kinds of difficulties outlined in the last section: the possibility of an agent's holding someone to account without believing there is any justification for doing so; and the possibility of an agent's holding someone to account without being at all susceptible to the reactive emotions. In contrast to Wallace's narrow approach, Strawson aims to elucidate the nature and role of the reactive emotions and their connection to demands or expectations in relation to broader moral and evaluative practice.<sup>19</sup>

For Strawson, to understand the practice of holding agents responsible we should start with the 'very great importance we attach to the attitudes and intentions towards us of other human beings, and the great extent to which our personal feelings and reactions depend upon, or involve, beliefs about these attitudes and intentions' (1962: 48). The reactive emotions are thus *reactions* to the perceived goodwill or indifference shown towards us by others as expressed in their intentions and attitudes (Strawson 1962: 49). As such, the reactive emotions express a demand for a certain degree of goodwill or regard from others towards ourselves.<sup>20</sup> For example, if a bus driver sees me running for the bus in the rain and drives away just as I reach the door, I am liable to feel resentment towards him in a way that I would not if I knew that he had not seen me. In resenting the bus driver in this way, I hold him accountable for his behaviour towards me and for the intentions and attitudes expressed by his behaviour. And holding him accountable in this way, on Strawson's model, is just to be prone to the reactive emotions (1962: 63)

---

<sup>19</sup> Strawson views the role of the reactive attitudes as central to any adequate account of the nature of morality and of moral responsibility: 'Only by attending to this range of attitudes can we recover from the facts as we know them a sense of what we mean, i.e. of all we mean, when...speaking the language of morals...' (Strawson 1962: 64).

<sup>20</sup> Negative reactive emotions, such as resentment, says Strawson, are 'precisely the correlates' of this demand or expectation 'where the demand is felt to be disregarded' (1962: 63).

These comments indicate that Strawson elucidates the reactive emotions or attitudes in relation to a general demand or expectation which derives from a fundamental concern that human agents have with how they are treated by others. This is at least the case with what Strawson terms ‘personal reactive attitudes,’ such as resentment and gratitude. But he stresses that human agents also make demands on themselves for others and demands on others for others (1962: 57). ‘Self-reactive attitudes’ such as guilt and shame characterise the first kind of demand<sup>21</sup>; whereas the second kind of demand is characterised by ‘generalised’ or ‘vicarious’ reactive attitudes such as moral indignation and moral approval.<sup>22</sup> Common to all these types of reactive attitudes is their dependence on the *normative assessment or evaluation* of the intentions, behaviour, and attitudes of those to whom they are directed. For instance, I resent the bus driver because he behaved in a needlessly indifferent and intentionally unhelpful manner towards me. These failings in his attitudes and behaviour I assess or judge make appropriate the attitude of resentment towards him. Similarly, I might feel moral indignation towards a shopkeeper who discriminates against customers on the basis of race, because I judge his racist attitude and behaviour to be morally wrong and indignation a warranted response to this wrong.

As we have seen, an agent may feel resentment towards someone without believing she has any grounds for such a feeling. The crucial feature of Strawson’s approach that allows us to account for such experiences is his broad reading of the mutual dependence claim. Implicit in his argument is the elucidation of the mutual dependence between the reactive emotions and the stance of holding to account at the level of general dispositions rather than at the level of specific cases of reactive emotion. On this model, to be broadly or generally susceptible to the range of

---

<sup>21</sup> For example, if I fail to do my fair share of the housework, I may violate a demand I make on myself to treat you fairly, and thus be liable to feel guilty as a result of recognising this violation.

<sup>22</sup> Strawson suggests these are feelings that stem from a demand for a degree of consideration from others towards all agents who are appropriate objects of moral concern (1962: 57). He appears to regard the generalised reactive attitudes as typically moral in character (1962: 56). But we need not follow him in characterising moral reactive attitudes in this way. Just as self-reactive attitudes of guilt and shame are often moral in character, generalised forms of disapproval or praise can be of a non-moral nature. For example, I may disapprove of the choice of curator because I think it will be bad for the gallery, without thinking that those who appointed the curator warrant moral condemnation.

reactive emotions is to be generally disposed to hold agents to account with regard to their behaviour; and to be such as to hold agents to account with regard to their behaviour is just to be generally susceptible to the reactive emotions (cf. Strawson 1962: 53, 54). To have such dispositions and susceptibilities, says Strawson, is constitutive of participation in normal interpersonal relations and in moral discourse and practice (1962: 52, 57-58).<sup>23</sup> It is against the background of this general orientation characteristic of the participant point of view that we can make sense of unwarranted reactive emotions and the failure to feel reactive emotions when they are judged warranted.

Consider first the case of unwarranted reactive emotions. An agent experiences an unwarranted reactive emotion when she experiences that emotion but does not believe it to be an appropriate response to the behaviour or act that elicits it. I have argued that because in such cases an agent fails to endorse her response, she cannot be said to have adopted the stance of holding someone to account, even though she experiences a genuine reactive emotion. Yet the mutual dependence claim tells us that reactive emotions are essentially connected to the stance of holding someone to account. The model of general susceptibility to the reactive emotions and sanctioning responses suggests a response to this problem. On that model, we understand what it is to experience genuine warranted *or* unwarranted reactive emotions in relation to the role the reactive emotions play in the broader practice of holding agents to account that characterises the participant point of view. An agent who believes a reactive emotion towards someone is unwarranted and who is generally disposed to hold others and herself to expectations will tend to discount that emotion and the kind of behaviour the emotion, when warranted, gives agents

---

<sup>23</sup> Strawson contrasts the 'participant' attitudes and their related practices with the range of 'objective' attitudes (1962: 52). We adopt the objective attitudes (appropriately) towards those who it is inappropriate to hold accountable for their behaviour because they are constitutionally incapable of acting responsibly: 'To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what... might be called treatment; as something certainly to be taken account...of; to be managed or handled or cured or trained; perhaps simply to be avoided...The objective attitude may be emotionally toned in many ways...: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. But it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other' (Strawson 1962: 52).



reason to engage in.<sup>24</sup> For example, an agent will not see the fact of his feeling unwarranted guilt as reason to apologise to anyone or to modify his behaviour in the way that he would were he to believe his guilt justified (cf. Rawls 1971: 482). Reactive emotions are thus made sense of in terms of central cases in which they are warranted responses to conformity with or breaches of expectations and in which they count as holding agents to expectations. These central cases characterise the general orientation of the participant point of view. Thus, a participant agent who experiences an unwarranted reactive emotion will still *be such* as to hold herself and others to expectations. On this model, then, the reactive emotions are essentially connected to the stance of holding agents to account, even though any particular experience of a reactive emotion may not count as holding someone to account in the absence of a relevant belief that the emotion is warranted.

The proposed model can also make sense of Wallace's claim that an agent can hold someone to account in cases where the agent fails to feel a particular reactive emotion she believes to be warranted. Consider again the example of the rude bus driver who intentionally drives off just as I reach the bus. Perhaps I have had a very tiring day and can't work up the resentment I think is warranted in such cases. It is crucial to note that my failure to feel resentment in this instance does not undermine the claim that I am disposed to resent people who are rude to me. Such psychological dispositions need not manifest invariably in order to be genuine dispositions. So while in this instance I fail to feel resentment I think is warranted, I am still disposed to feel resentment in response to such situations.

Consider how my proposed model of moral judgement would work in relation to this example. I judge that the bus driver has acted wrongly because he failed to treat me with appropriate consideration and respect. Presumably I believe there is a moral requirement that prescribes treating others with due consideration and respect in such contexts and I believe that the bus driver has violated this requirement. Let's focus, then, on the judgement that the bus driver is morally required to treat her passengers

---

<sup>24</sup> For example, an agent will not see the fact of his feeling unwarranted guilt as reason to apologise to anyone or to modify his behaviour in the way that he would were he to believe his guilt justified (cf. Rawls 1971: 482).



with due consideration and respect. What conditions must be met in order for this judgement to be sincere? The sincerity condition for the assertive portion of the judgement is that I believe that the bus driver is morally required to treat her passengers with due consideration and respect. The sincerity condition for the commissive portion of the judgement is that I normatively expect the bus driver to conform to the requirement to treat her passengers with due consideration and respect. To hold someone to a requirement in this way is to have a certain belief and to have certain dispositions. If I normatively expect the bus driver to conform to this requirement then I must *believe* that a violation of the requirement would warrant a negative sanction and I must *be disposed to sanction* negatively violations of this requirement. These conditions are summarised below:

**Moral Judgement:** S judges that the bus driver is morally required to treat her passengers with due consideration and respect.

**Sincerity Conditions:**

1. S **believes** that the bus driver is morally required to treat her passengers with due consideration and respect.
2. S **normatively expects** the bus driver to conform to the requirement that she treat her passengers with due consideration and respect:
  - a) S believes that violations of this requirement warrant negative sanctions.
  - b) S is disposed to sanction negatively violations of this requirement.

In this chapter I have been trying to understand what it is for an agent to be sincere in her commitment to a moral requirement. My claim is that the posited sincerity condition for normative expectation captures the robust sense of endorsement or commitment that, according to moral judgement internalism, is necessarily tied to moral judgements. We start with the internalist intuition that there is something more to moral judgement than bare belief: something involving a notion of endorsement or commendation that is significantly tied up with motivation. This notion of endorsement is connected to the commissive aspect of the moral speech act. When we judge that something is morally right we commit ourselves to being guided by the content of the requirement expressed by our judgement. We are

sincere in this commitment when we hold ourselves and others to account with regard to the relevant requirement. The notion of holding to account or normative expectation is intimately tied up with the kinds of reactive emotions that typically characterise moral practice. My claim is that something like this notion of normative expectation plays a central role in our moral psychology. It is what is missing from both purely cognitivist accounts of moral judgement and traditional non-cognitivist accounts of moral judgement. Purely cognitivist accounts of moral judgement fail to account for the endorsing function of such judgements. Traditional non-cognitivist accounts of moral judgement fail to account for the distinction between merely desiring or being motivated to pursue an option and endorsing the pursuit of an option (see chapter 1.6).

To be committed to something is, *inter alia*, to be disposed to promote or support that thing. One cannot be committed to something if one has absolutely no inclination to promote or support it. So my account of sincere commitment to a moral requirement is designed to capture the sense in which we are inclined or disposed to support and promote such requirements. In this section, I have argued that the model of general susceptibility to the reactive emotions can accommodate a version of the mutual dependence claim while accounting for the problem of unwarranted reactive emotions and the problem of holding someone to account in the absence of specific reactive emotions. In chapter three, I said that a preparatory condition for the speech act of moral evaluation is that the speaker is normatively engaged. We are now in a position to give a more precise account of what this means. For someone to be normatively engaged is, on my account, for her to be generally susceptible to the reactive emotions and sanctioning responses which in turn enables her to hold agents to expectations and thus to engage in the normative assessment of those agents' intentions, attitudes, and behaviour.

## 7. Moral Judgement, Normative Competence and the Amoralist

One of the main themes of Stocker's paper is that motivation and evaluation do not have a simple and direct relation to each other: 'Rather, they are interrelated in various and complex ways, and their interrelations are mediated by large arrays of complex psychic structures, such as mood, energy, and interest' (1979: 739). This is why he says that when agents are depressed, or lack interest or energy, they often fail to intend to do what they believe it is best to do – the standard case of *akrasia* or weakness of will. The discussion thus far in this chapter supports both the contention that the relation between motivation and evaluation is not simple and direct and the claim that agents often fail to intend to do what they believe it is best to do. However, the discussion has also been an attempt to explain the view that moral judgements are of conceptual necessity tied to endorsing attitudes which play a central role in moral motivation: the version of moral judgement internalism defended in chapter 1.

There is not a direct conflict between the phenomenon of weakness of will and simple moral judgement internalism. An agent suffering from weakness of will only fails to be sufficiently motivated to perform the act she judges it is best to do; and this is consistent with her having some motivation to perform that act and thus with there being a necessary connection between the judgement that the act is right and motivation to perform the act. Nonetheless, opponents of moral judgement internalism often raise the possibility of cases in which an agent not only fails to intend to do what she judges is best, but fails to have any motivation inclining her towards her better judgement. And the possibility of this agent does seem to threaten the thesis of simple moral judgement internalism. In chapter 1. I argued that we should reject the simple version of moral judgement internalism. Part of the reason for this is that simple moral judgement internalism neglects the complex nature of the relation between evaluation and motivation highlighted by Stocker. I said in 1.6 that simple moral judgement internalism is in one sense too strong and in another sense too weak. It is too strong because agents sometimes have no desire to perform acts they judge they are morally required to perform. It is too weak because merely

having a desire to perform an act one believes to be right is not sufficient to count as endorsing that act or the requirement that prescribes it.

The version of moral judgment internalism I defended in chapter 1 asserts a connection of conceptual necessity between the act of judging that an option is morally right and endorsing or commending that option. In this chapter I have argued that in order to make a genuine moral judgement an agent must be affectively or conatively disposed to be guided by the content of her moral deliberations. This means that genuine moral judgements presuppose a background of affective-conative engagement with a particular kind of normative practice. There is thus a non-defeasible conceptual or 'internal' connection between genuine moral judgements and the range of conative attitudes necessary for normative engagement. This position is clearly distinct from the externalist view that the possession of a relevant moral belief is sufficient for an agent to make a genuine moral judgement. So there is a kind of internalism at the level of the preparatory conditions for genuine moral judgements that distinguishes my view from moral judgement externalism.

This form of internalism at the level of preparatory conditions helps to explain the more standard internalist claim that moral judgements necessarily express endorsing or commendatory attitudes. So my position is also distinguished from externalism in relation to the kinds of attitudes it says moral judgements necessarily express. Specifically, my view says that moral judgements express an attitude or stance of normative expectation in addition to a moral belief. The stance of normative expectation comprises a belief that violations of the requirement expressed by the moral judgement warrant negative sanctions and a disposition to sanction negatively violations of this requirement. So if an agent judges that she is morally required to  $\Phi$ , then she is disposed to sanction negatively violations of the requirement to  $\Phi$ . The notion of normative expectation is meant to capture the strong sense of commitment or endorsement associated with moral judgements. This too distinguishes it from moral judgement externalism.

One of my central claims, then, is that if we reflect on the concept of endorsement involved in moral judgement it appears to involve something close to the stance of

normative expectation. Something like this stance seems to be needed to account for failures of motivation and, crucially, for the distinction between merely desiring an option and endorsing an option. As I mention in chapter 5.1, the notion of normative expectation also seems consistent with recent speculation about the nature of ‘higher cognitive attitudes,’ such as guilt, resentment and moral indignation. Research into such attitudes suggests that they are highly integrated with cognitive activity involved with long-term actions, coordination and planning (Griffith 1997: 100). It is plausible to suppose that a stance like that of normative expectation is involved in the kind of normatively governed intentional behaviour we associate with human agents (cf. Gibbard 1989: chap. 4; and see 5.1 above).

In order to understand the different kinds of failures of moral motivation and their implications for moral judgement internalism, I want to reflect on the complex picture of moral psychology and moral motivation outlined in previous chapters. I mentioned in chapters one and two that different notions of motivation seem to figure in moral deliberation at a number of different levels. Sometimes talk of ‘motivation’ is talk of motivating states: some psychological states are considered essentially or internally motivating in the sense that to be in such a state constitutes being motivated. Typically, we talk of desires and wants in this respect, but there are many other *conative* states that are broadly motivating in this same sense, including such attitudes as sympathy, anger, disapproval and concern. The non-normative motivating state sense of ‘motivation’ should not be confused with the sense in which someone may be motivated by a particular consideration she takes to be justifying.<sup>25</sup> For example, the consideration or reason in light of which I helped her was that she needed assistance. Though I felt sympathy for her, I did not take the fact of my feeling sympathy to be a consideration in favour of helping her.<sup>26</sup> The sympathy I feel towards her is, in this case, a non-normative motivating attitude. The consideration or reason in light of which I act is also to be distinguished from a belief with a justifying reason claim as its content, such as my belief that she needs

---

<sup>25</sup> See 5.3 for an outline of the notion of a favouring consideration.

<sup>26</sup> Though the fact that she warranted sympathy, if she did, may have been part of my justification for helping her.

assistance. In this case, it is the consideration that she needs assistance and not the belief that she needs assistance that I take to be justification for my helping her.

Now, it is clear that the model of moral judgement internalism I have put forward does not posit a necessary connection between making a moral judgement and being motivated by the justifying content of that judgement. That is, it does not posit a necessary connection between moral judgement and what we might call *normative motivation*. And this accords with what Stocker has to say about the influence of mood, interest, and energy on our moral motivation. John's depression means that he is not motivated by the consideration that he has a moral obligation to vote in the national election: a consideration he takes to be justifying. However, if John is sincere, he must, in addition to believing his moral judgement, normatively expect himself to act in accordance with the proposition that it is right for him (and all relevant others) to vote in the national election. And, as we have seen, in order to hold himself and others to an expectation in this way, John must be disposed to sanction violations of the requirement expressed by his judgement, and to do this he must be generally susceptible to the reactive attitudes and sanctioning responses that characterise the participant point of view.

If the point of moral practice and moral judgement is to guide conduct, then we would expect the reactive attitudes to help facilitate this aim. And it seems that they do. The reactive attitudes put pressure on action: they help to reinforce the aim of getting agents to be motivated by normative judgements. When the match between normative judgement and motivation breaks apart, as in the case of weakness of will and mood-based motivational disorders, reactive attitudes help to bring motivations back in line with normative judgements. Reactive attitudes also exert interpersonal pressure on agents to coordinate their normative judgements with their motivations.

Given the interpretation of moral judgement internalism advocated above, what can we say about the issue of amoralism and of failures of moral motivation more generally? This will depend on whether our concern is with amoralist scepticism or



with unprincipled amorality.<sup>27</sup> Recall that the amoralist sceptic is someone who questions the practical justification of moral demands, whereas the unprincipled amoralist is someone who accepts the justification of moral demands but who fails to be motivated by them. Given the above distinction between normative and non-normative motivation, we can distinguish two possible forms of unprincipled amorality. One form of unprincipled amorality involves the agent who judges that an act is morally required but who fails to be *normatively* motivated by her judgement. Another form of unprincipled amorality involves the agent who judges that an act is morally required but who fails to have *any* conative attitudes inclining her towards the act. On the model I have proposed, the latter kind of unprincipled amoralist will not count as having made a *sincere* moral judgement because she lacks general susceptibility to those attitudes that would allow her to hold agents to expectations. More fundamentally, her lack of any susceptibility to these attitudes prevents her from being normatively engaged in the way required to meet the precondition for making *genuine* moral judgements. This means I am committed to denying the possibility of this kind of unprincipled amorality. By contrast, the unprincipled amoralist who fails to be normatively motivated by her moral judgement is potentially capable of making both genuine and sincere moral judgements. This is the familiar example of the depressed or mood-affected agent outlined by Stocker. Let's call the unprincipled amoralist who makes a genuine moral judgement but who fails to be normatively motivated by it the 'incontinent amoralist'.

The incontinent amoralist can make genuine and sincere moral judgements because she can be generally susceptible to the reactive attitudes and thus generally disposed to hold herself and others to expectations, even though she fails to be normatively motivated by her moral judgements. So the possibility of incontinent amorality is consistent with there being a necessary connection between moral judgements, particular dispositions to sanction requirements, and the range of conative attitudes that characterise the participant point of view. In fact, a consequence of the position I have proposed is that the possibility of incontinent amorality actually *depends* on

---

<sup>27</sup> See chapter 2, section 3 for a discussion of these different forms of amorality.

the incontinent amoralist's being normatively engaged in this way. This is because it is only if the incontinent amoralist is generally susceptible to the reactive attitudes that she is capable of holding herself and others to account in the way necessary to make genuine moral judgements. So, on my model, rather than confuting moral judgement internalism, incontinent amoralism actually entails it.

The incontinent amoralist can fail to have sufficient normative motivation to do what she judges is right or she can fail to have any normative motivation to do what she judges is right. In either case, incontinent amoralism is consistent with my version of moral judgement internalism. But perhaps the proponent of anti-internalist amoralism has something more extreme in mind than the incontinent amoralist. This is the other type of unprincipled amoralism: the view that there are agents who make genuine moral judgements in the absence of any conative attitudes inclining them towards the acts favoured by their judgements. Now, as I have said, the position I have adopted commits me to denying the possibility of this form of amoralism. The plausibility of this denial will, of course, in part depend on the success of the model of moral judgement I have defended and the success of that model in accounting for a range of failures of moral motivation. The hope is that because my model *can* account for various examples of such motivational failure, it can capture the core of the intuitions that generate the arguments from amoralism, even while it excludes extreme examples of unprincipled amoralism.

If we reflect on my account of the nature and structure of moral judgement, I think we can begin to appreciate the force of the intuitions that deny that the extreme unprincipled amoralist makes genuine moral judgements. On my model, an agent's moral judgement is assessed as sincere or not in relation to whether she believes the relevant moral proposition and whether she holds herself and others to account with regard to that proposition: both conditions must be met in order for the moral judgement to be sincere. It is worth noting, however, that the mental states involved in moral judgement do not make up a compound attitude comprising entirely

unrelated states.<sup>28</sup> We need to understand the role of moral judgements in relation to broader moral and evaluative practice. The point of such a practice is to guide conduct. This is reflected in the illocutionary point or essential condition of moral speech acts (see 3.6). There are a range of conceptual connections between the different states involved in moral judgement that allow them to function together to achieve the aim of guiding conduct.

On the model I have presented, a moral judgement involves a belief about a moral requirement, a belief about the kinds of responses a violation of that requirement would make appropriate, and a tendency to sanction such violations. A failure of moral belief will bring with it a failure of normative expectation. This is because normative expectation requires both a disposition to sanction and a belief that the sanction is warranted, and if the agent does not believe there is a genuine moral requirement, then she cannot consistently believe that a violation of that ‘requirement’ would warrant some form of negative sanction. This reflects the conceptual dependence of the reactive emotions involved in normative expectation on normative beliefs about the propriety of such emotions. So while it is sufficient to undermine the sincerity of a moral judgement that an agent either fails to believe the relevant moral proposition or fails to hold herself and others to account with regard to it, a failure of the belief portion of the sincerity condition for moral judgements will tend to bring with it a failure of the normative expectation portion of the sincerity condition for moral judgements.

For an agent to be normatively engaged is for her to be capable of and disposed to hold herself to requirements. This means that insofar as an agent believes some requirement warranted, there will be, given her engaged stance, a range of reactive conative attitudes inclining her towards that requirement. If the functioning of the reactive attitudes that allow an agent to act from the participant point of view is so

---

<sup>28</sup> Mele seems to have this kind of compound attitude in mind with his notion of an MR belief\* (1996: 740). An MR belief\* is any belief\* that one is oneself morally required to do A, where belief\* is a compound of two distinct and unrelated attitudes, one essentially truth-seeking (a standard belief) and one essentially motivational (a standard desire). Mele dismisses the notion of this kind of compound attitude. See my section 8 above for a discussion of Mele’s conception of moral judgement internalism.

severely inhibited that she no longer has any tendency or capacity to hold herself to account, the agent will no longer count as being normatively engaged, and so will no longer be able to make genuine moral judgements: she fails to meet what I called the preparatory condition for moral judgements. As an example, consider again the case of John's failure to vote in the national election even though he believes he is morally required to do so. If John makes a genuine moral judgement, he must hold himself and others to the moral requirement despite his failure to be normatively motivated by it. Perhaps prior to the election he feels and expresses approval towards those who intend to vote, and after the election he disapproves of those who failed to vote and feels guilt himself for his failure to vote. But if John's depression is so severe that he is simply unable to feel anything about those things he believes to be right, then, on my approach, he is incapable of holding himself and others to requirements and incapable of making genuine moral judgements.

We should not confuse the agent whose mental state is such that he is incapable of any reactive emotional response with the agent who temporarily fails to manifest appropriate reactive emotions. Jane believes she is morally required to vote but fails to do so because she has just fallen in love and spends the day of the election, and much of the next week, with her lover. Jane does not forget about the election, but is completely preoccupied with her current situation. Even during the next week, she cannot seem to think of anything or anyone but her new lover; she certainly does not have the emotional room to feel guilty for not voting. Nonetheless, Jane knows that her failure to vote warrants feeling guilty, and as her passions cool and the consequences of her failure to vote in a marginal electorate become clear – the election of a government to whom she is ideologically opposed and the institution by them of policies she finds objectionable – she begins to feel bad about having not voted. The complex network of dispositions and attitudes that make up the participant point of view – the general tendency to hold herself and others to account – was always intact in Jane, and throughout the entire incident she held herself to the expectation that she vote in the national election. By contrast, John is so severely affected by depression that his *general* susceptibility to the reactive attitudes is undermined and he becomes incapable in general of holding himself or anyone else

to account. We might say that the unity of John's moral identity has been fragmented to the point where it is no longer reasonable to attribute to him a coherent moral point of view.

It should now begin to be clear why determining the status of moral judgements can be both difficult and contentious. Determining whether or not to ascribe a moral judgement to an agent depends in part upon an assessment of the competence of the agent to make such judgements. Only if an agent is normatively competent is she capable of holding herself and others to account in the way necessary to make genuine moral judgements. In general, normative expectation will track moral belief in the normatively engaged agent; insofar as it does not, we will tend to consider the agent as lacking normative competence and withdraw ascriptions of moral judgement. The central idea here is that we have presuppositions about the kind of agents who are capable of making genuine moral judgements: who are capable of being moral appraisers. What I have been considering, following Strawson, is that the capacity to hold agents to account and the role of the reactive emotions in this capacity are an integral part of our ordinary conception of moral competence. More than this, however, I think that the capacities of the normatively engaged agent are an integral part of our ordinary notion of the morally responsible agent.<sup>29</sup> Below, I explore some connections between the kind of normative competence necessary for making moral judgements and the kind of moral accountability necessary to be an appropriate object of moral appraisal. I relate these issues to the problem of the extreme amoralist.

My suggestion is that the extreme amoralist suffers from a degree of normative incompetence that places her outside ordinary moral practice and which makes it inappropriate to ascribe moral judgements to her.<sup>30</sup> To support this view, I want to

---

<sup>29</sup> Cf. Wallace 1994: 53: 'To view people as responsible... is to see them as autonomous agents who are reflective about their lives, who have a set of values and commitments sufficiently structured to constitute... a conception of the good, and who aim to advance that conception in their action.'

<sup>30</sup> Such normative incompetence does not mean that the agent in question is not an appropriate object of moral *concern*. Indeed, we may acquire specific moral obligations towards such an agent precisely because of her status.



consider some conditions under which we tend to exempt agents from moral responsibility. Sometimes we judge it inappropriate to hold agents morally responsible, not because in particular instances they are blameless, but because they are in general inappropriate objects of this kind of moral appraisal (cf. Strawson 1962: 50-52, 58-59). As Strawson points out, the conditions under which we exempt an agent from general responsibility 'do not invite us to see the agent's action in a way consistent with the full retention of ordinary interpersonal attitudes and merely inconsistent with one particular attitude' (1962: 51). Such agents are typically 'psychologically abnormal' or 'morally undeveloped', and it is thus appropriate to adopt towards them the objective range of attitudes in place of the reactive attitudes (Strawson 1962: 52).<sup>31</sup> This seems correct. For example, we do not generally hold very young children responsible for their behaviour; and this is in large part due to their lack of cognitive and emotional development. Adopting the objective stance towards such children means training and managing their behaviour rather than apportioning blame.<sup>32</sup> It also means that we do not ascribe moral judgements to them. And this is not simply because they are unable to grasp moral concepts. Such children lack the capacity to regulate their conduct in a way that would make it appropriate to appraise them morally. A large part of this lack of capacity is their inability to hold themselves and others to expectations: an inability that also means they cannot make genuine moral judgements. So one of the factors that makes it inappropriate to hold young children morally responsible also makes it inappropriate to ascribe moral judgements to them.

Strawson's model suggests that we tend to take up the objective stance towards those who lack the reactive attitudes or whose reactive emotional capacity is inhibited in some serious way. Such agents suffer from a kind of deep practical normative

---

<sup>31</sup> See note 22, above, for the distinction between the objective and the reactive range of attitudes. Strawson notes that we sometimes adopt the objective attitudes towards normal responsible agents for purposes of therapy or 'as a refuge... from the strains of personal involvement' (1962: 52).

<sup>32</sup> Of course, a crucial part of this training will be the rehearsal of disapprobation and approval towards the child to bring her towards moral competence. The case of children illustrates why ascriptions of moral judgement and of moral responsibility can be difficult: 'Thus parents and others concerned with the upbringing of young children cannot have to their charges either kind of attitude [reactive or objective] in a pure or unqualified form. They are dealing with creatures who are potentially and increasingly capable both of holding, and being objects of, the full range of human and moral attitudes, but are not yet truly capable of either' (Strawson 1962: 61-62).



incompetence or immaturity. My claim is that the objective stance is the appropriate stance to take towards the extreme amoralist as he too suffers from this kind of practical normative incompetence. Consider one of Brink's examples of amorality, the sociopath (1989: 46). The sociopath would seem to be a clear example of someone who is outside the kind of normative practice characterised by susceptibility to the reactive attitudes. The following comments by Strawson capture this point:

But suppose we see the agent...as one whose picture of the world is an insane delusion; or as one whose behaviour...is unintelligible to us, perhaps even to him, in terms of conscious purposes, and intelligible only in terms of unconscious purposes; or even, perhaps, as one wholly impervious to the...reactive attitudes...wholly lacking...in a moral sense...[T]o the extent to which the agent is seen in this light, he is not seen as one on whom demands and expectations lie in that particular way in which we think of them as lying when we speak of moral obligation; he is not, to that extent, seen as a morally responsible agent, as a term of moral relationships, as a member of the moral community. (Strawson 1962: 59)<sup>33</sup>

By adopting the objective attitudes towards the sociopath, we signal that it is inappropriate to treat him either as an object of moral appraisal or as a subject capable of moral appraisal. Because he is unable to hold himself or others to account, the extreme amoralist is incapable of being morally responsible for his actions and is incapable of making genuine moral judgements.<sup>34</sup>

## 8. Bare Moral Belief

In the previous section, I rejected the extreme form of amorality that is said to undermine moral judgement internalism. However, it needs to be made clear precisely what has been denied, as some opponents of moral judgement internalism will undoubtedly object that I have failed to capture the real force of the externalist's amoralist argument. While I have shown that the extreme amoralist cannot make a genuine moral judgement, construed as a moral speech act, the externalist might argue that moral judgement externalism, in its most potent and interesting form, is a

---

<sup>33</sup> Watson (1987) discusses the issue of moral responsibility and the sociopath/psychopath in relation to Strawson's views on the participant attitudes and moral community.

<sup>34</sup> Insofar as such an agent does warrant moral indignation or blame, we must hold that he is an appropriate object of moral appraisal and, as such, capable of holding himself and others to expectations. But the extreme amoralist is not someone whose evaluative judgements are merely horribly misguided – someone, for example, who believes he is justified in brutalising women and normatively expects himself to do so; he is someone who lacks any, or sufficient, susceptibility to the kinds of attitudes that would allow him to hold himself and others to account.

thesis about moral psychology rather than moral language. And nothing I have said shows that it is logically impossible for the extreme amoralist to *believe* that he is morally required to perform some act and have no motivation to act in accordance with that requirement.

We saw in chapter one that much of the difficulty in assessing moral judgement internalism stems from a lack of clarity about the precise nature of the thesis under consideration. I argued that there are good reasons for interpreting moral judgement internalism as a thesis about judgement acts rather than judgement contents. Some theorists seem to view moral judgement internalism as equivalent to what we might call moral belief internalism.<sup>35</sup> I said in chapter one that Dancy interprets Nagel as holding a view of internalism along these lines. Dancy interprets moral judgement internalism as the doctrine that moral judgements express beliefs which ‘cannot be present without motivating’ (1993: 3).<sup>36</sup> Like Dancy, Alfred Mele interprets Nagel’s version of internalism as a form of moral belief internalism (1996: 729).

Specifically, he argues that in order to capture a ‘robust’ version of internalism in which motivation is ‘built into’ moral judgements, moral judgement internalism should be construed as, what he terms, ‘constitutive internalism’.<sup>37</sup>

Necessarily, any belief that one is (oneself) morally required to A constitutes motivation to A. (Mele 1996: 731)<sup>38</sup>

Mele is persuaded, again like Dancy, that Nagel’s talk of motivation being ‘guaranteed by the truth of ethical propositions’ should not be interpreted as a form

---

<sup>35</sup> For example, Dancy characterises moral cognitivists as holding that ‘moral judgement is belief’ (1993: 7)

<sup>36</sup> Because of the problems such a position has accounting for mood-based failures of motivation or ‘accidie’, Dancy rejects strict entailment between moral belief and motivation. He defends a purely cognitivist account of moral motivation according to which moral beliefs are ‘intrinsically motivating states, which can be present without motivating but which when they do motivate do so in their own right’ (1993: 24).

<sup>37</sup> The constitutive reading of internalism is intended to block the possibility of a necessary connection between moral belief and motivation being explained by factors external to the moral belief: for example, by desire or motivation of some sort being a necessary precondition for coming to hold genuine moral beliefs (Mele 1996: 730).

<sup>38</sup> Mele notes that he has ‘no objection to substituting ‘judgement’ for ‘belief’ (1996: 731). He limits his account of internalism to first person moral ought beliefs; my objections to moral belief internalism are not affected by this restriction.

of content internalism about motivation (1996: 729).<sup>39</sup> However, as I have argued, Nagel does not assert a necessary connection between moral judgements or beliefs and motivating states but a necessary connection between moral requirements and *justifying* motives for action (and thus a necessary connection between judging that something is a moral requirement and judging that it provides a justifying motive for action). This is the position I called content internalism about justifying reasons. Mele's reading of Nagel leads him to argue for the thesis that internalism requires any genuine moral belief to *constitute* a motivating *state* (1996: 730). And this is the position he takes to be ultimately untenable because, not only is it contradicted by instances of unprincipled amorality arising from 'listlessness' or accidie, but it depends upon the dubious notion of 'a noncompound attitude that is – at once and essentially – receptive to a species of moral truth and constitutive of motivation to engage in suitable behaviour' (Mele 1996: 740).<sup>40</sup>

I agree with Mele that the notion of a noncompound essentially motivation constituting and truth-seeking state is suspect. However, a state of this kind is not entailed by Nagel's version of internalism and is not something which moral judgement internalists are, or should be, required to accept. Why not? Well, it presents moral motivation as a form of state-based non-normative motivation. The problem with this view is that if moral considerations play a genuine role in deliberation about what to do, then it seems it must be possible that agents are sometimes motivated by the normative content of their moral judgements.<sup>41</sup> The state-based view misses the sense in which moral motivation is a species of normative motivation. Rather than supporting the state-based view, Nagel's approach supports a conception of moral judgements as having essentially justifying content capable of motivating moral agents (1970: 65). On this model, it makes

---

<sup>39</sup> Mele refers to what I have called content internalism about motivation as 'requirement internalism' (1996: 729). Recall that this is the view that an agent's having a motivation to perform or promote some act is part of what it is for a moral proposition requiring her to perform that act to be true (see chap 1.3)

<sup>40</sup> Mele argues that were constitutive internalists to discover that there were no such things as intrinsically motivating ought beliefs it would be more likely they would modify their conception of the nature of moral motivation than conclude that no one has ever had a genuine moral belief or has ever had a moral obligation: 'This is an indication that morality does not conceptually or metaphysically require that moral ought beliefs of the kind at issue be in the psychological repertoire of moral agents...' (1996: 751-752).

sense that the attitude agents have towards moral propositions is belief-like rather than desire-like.

Another consideration that counts against construing moral judgement internalism as a form of belief internalism is that doing so prevents traditional forms of non-cognitivism from being counted as internalist. Non-cognitivists cannot assert a necessary connection between moral beliefs and motivation because they deny that moral judgements express beliefs. However, this sits awkwardly with the fact that non-cognitivism is normally advanced as an exemplar of moral judgement internalism.<sup>42</sup> And I think we should be sceptical of any formulation of a familiar doctrine which excludes one of the standard exemplars of that doctrine. One of the central points at issue in the internalism debate is the nature of the psychological state or states involved in moral assent. In assuming the state involved is belief, moral belief internalism is simply question-begging with regard to this issue.

The opponent of the speech act approach to moral judgements may accept that moral judgement internalism should be characterised in a way that is neutral between different mental states. Nonetheless, she is likely to stress that given the speech act account is cognitivist, it does not block the possibility of an agent's having a bare moral belief independently of any motivations to act in accordance with the requirement expressed by the belief. Now, I am not inclined to deny the *logical* possibility of bare moral beliefs of this kind. What I do want to deny is that someone who fails to be normatively engaged, and who is thus unable to hold themselves or others morally responsible, is capable of making genuine moral evaluations. The crucial point to make is that the conception of moral judgement I have defended, which incorporates the notion of normative expectation and a particular conception of normative competence, is closer to our ordinary notion of what it is to make a moral judgement than is the concept of a moral belief possessed by the extreme amoralist: someone who, though devoid of moral emotions, is capable of representing moral facts. Part of the general intuition underlying internalism is that

---

<sup>41</sup> cf. chap. 1.5., above.

<sup>42</sup> Non-cognitivism's ability to accommodate internalism is often put forward as a reason to accept it. See chap. 1.2 for a discussion of this aspect of non-cognitivism.

in making a moral judgement an agent is *doing* something that goes beyond mere cognition. Part of what she is doing, I have suggested, is holding agents morally accountable for their behaviour: something that cannot be done merely by believing that some fact obtains.

Even if we construe moral judgement internalism as a thesis about psychological assent to moral propositions, it is still open to the speech act advocate to explain the apparently dual nature of moral assent in terms of the linguistic rules that are part of evaluative practice. The linguistic rules encode aspects of the practice – its point and presuppositions – and in turn influence and are an integral part of the practice. It is part of moral practice in general that moral assent involves belief and normative expectation. Insofar as agents are part of this practice, insofar as they are moral agents, they must follow the rules constitutive of the practice. These points are quite general and need to be expanded upon. I take up this task in the next chapter.

## **9. Conclusion**

In this chapter, I have argued against intention as the sincerity condition for the commissive portion of moral judgements because it cannot deal with the kinds of failures of motivation that result from depression and other mood-based disorders. By contrast, the stance of holding someone to account, or normative expectation, allows that an agent can be sincerely committed to a moral requirement even if, due to depression or fatigue, she fails to be normatively motivated to act in accordance with the requirement. We have seen that to be capable of holding someone to account an agent must be generally susceptible to the reactive emotions and sanctioning responses that characterise the participant point of view, and that to be susceptible in this way is just to be normatively engaged. On the proposed model, then, there is a necessary connection between making a genuine moral judgement and being susceptible to a range of conative attitudes. It is in holding herself and others to account in making a moral judgement that an agent's judgement goes beyond mere belief or cognition. And this is why the extreme amoralist cannot make genuine moral judgements: she cannot hold agents to account because she is not subject to the reactive emotions. The proposed model can thus accommodate various

types of moral motivational failure, and associated versions of amoralism, while supporting intuitions that exclude those who suffer severe forms of practical normative incompetence from making genuine moral judgements.



---

## **Moral Thought, Intentional Action and the Desire/Belief Thesis**

### **1. Introduction**

I begin this chapter with a discussion of the connection between moral thought and moral language on the proposed account. This is necessary if I am to defend the notion that the speech act model has implications for the nature of moral psychology. The analysis of moral speech acts suggests that moral motivation is more complex than is often claimed. Motivating states figure in the endorsing attitudes entailed by agents' moral judgements and motivation figures, in some sense, in agents' responses to the content of their moral judgements. I argue that we can distinguish two senses of the term 'motivation' corresponding to a distinction between motivating states and (being motivated by) motivating considerations. The distinction is used to address the issue of whether moral beliefs are capable of motivating independently of desires and the status of the desire/belief model of motivation and action explanation.

### **2. Moral Thought and Moral Language**

I have argued throughout this work that we need to take seriously the non-cognitivist claim that moral judgements in some sense go beyond mere cognition or description. The speech act approach is thus in part an attempt to accommodate the endorsing, commending or favouring aspect that seems central to ordinary notions of moral judgement. Unlike non-cognitivism, however, the speech act approach holds that moral judgements can be both commendatory and assertive in nature. Non-cognitivism traditionally concerns itself with an analysis of moral language, and draws its conclusions about moral psychology from this analysis. I think that this emphasis on moral language is largely correct: moral practice is an overwhelmingly social and public enterprise that predominately takes place via linguistic interchange. In this section, I suggest, following Gibbard, that moral practice is 'linguistically infused' in a way that makes the study of the structure of moral language and moral

speech acts particularly useful in coming to understand the nature of moral judgements (Gibbard 1990).

Non-cognitivists argue that any adequate analysis of moral judgements must capture those judgements commendatory or endorsing function. The notion of holding someone to a requirement or expectation captures, I think, a robust sense of endorsement or commendation. In general, if an agent endorses a requirement, we would expect her to be disposed to sanction positively conformity with that requirement and to sanction negatively violations of that requirement. Compare this view of endorsement with the view that the mere presence of a motivating state secures the endorsing function of moral judgements. Standard formulations of moral judgement internalism require that if an agent judges that it is right for her to  $\Phi$  in circumstances  $C$ , then she is motivated to  $\Phi$  in  $C$  (e.g. Smith 1994: 61). Smith argues that this is a very strong version of internalism, but in many ways it expresses quite a weak motivational requirement. It simply insists that the agent has some motivation to perform or promote the act she believes to be right; it does not require the agent to be motivated by the perceived rightness of the act or for her to in any way endorse the act. For instance, suppose that John believes it is right to give money to charity, but fails to have any desire to do so until someone convinces him that doing so would make others think better of him. Now John believes it is right to give money to charity and desires to give money to charity, so satisfying the internalist requirement. But the presence of *this* desire seems irrelevant to the genuineness of John's moral stance: it is not an expression of his endorsement of the moral requirement that one should give money to charity. While it is true that internalism states a necessary and not a sufficient condition for genuine moral judgements, if it is not to be empty as a claim about morality it must state a motivational condition that is clearly related to the status of moral judgements. Moral judgements go beyond belief in involving endorsing attitudes, but they also go beyond desire. For example, an agent can desire to eat the rest of a birthday cake without holding herself to a requirement that she eat the cake; and she can hold herself to the requirement that she not eat the rest of the cake, while desiring very much to eat all the cake. To desire to perform an act is not, it would seem, the same

as to commend or endorse the performance of the act.<sup>1</sup> I take it that any adequate account of moral judgement needs to make sense of this distinction.

Gibbard in his version of non-cognitivism attempts to capture the intuitive distinction between merely desiring to perform an act and endorsing the performance of an act (1990: chap. 4). To this end, he distinguishes between motivation characterised by the *acceptance* of a norm or requirement and motivation characterised by desires that arise from basic impulses such as hunger and thirst (Gibbard 1990: 56). He also distinguishes between motivation characterised by having merely ‘internalised’ a norm and motivation characterised by acceptance of a norm (Gibbard 1990: 68-71). For instance, an agent may have internalised a norm or requirement to queue patiently in shops, but accept that in this instance, given that someone has pushed in, it is appropriate to make a fuss. Often in such cases the agent will not do what he thinks he should do because, in Gibbard’s terms, the agent is ‘in the grip of a norm’. In this example, the agent is in the grip of a norm prescribing patient queuing, even though he accepts that what he should do is take the queue-jumper to task over her behaviour. We would say that while the agent endorses or commends telling the queue-jumper to return to the back of the queue, his overriding motivation is to stay quiet and wait patiently to be served.

Ultimately, Gibbard characterises the notion of genuinely accepting a norm in terms of an agent’s tendency to avow the norm publicly in unconstrained contexts (1990: 74).<sup>2</sup> As Gibbard’s view is non-cognitivist, he cannot invoke *belief* in the validity of a norm in order to capture the notion of genuine acceptance. I think this makes his account of norm acceptance excessively behaviouristic, but I do not want to dwell on this aspect of his theory. Gibbard aims to provide an account of norm acceptance in evolutionary terms. He argues that the notion of normative avowal and the kind of

---

<sup>1</sup> Cf. Korsgaard (1997: 234): ‘In Kant’s view, an inclination is a kind of attraction to something, which is grounded in our sensuous nature, and in the face of which we are passive. By themselves, inclinations have no normative force; they are not reasons.’ Presumably we might endorse some desires or inclinations if we judge we have reasons for holding them (see 5.4, below).

<sup>2</sup> He also stresses the role of demands for consistency in normative discourse in contributing to the acceptance of norms: ‘To accept a norm... is in part to be disposed to avow it in unconstrained normative discussion, as a result of the workings of demands for consistency in the positions one takes in normative discussion’ (Gibbard 1990: 74).

public discussion that characterises taking up and avowing normative positions – what he terms ‘normative discussion’ – can plausibly be viewed as fitness-enhancing forms of behaviour in humans because they facilitate improved social coordination and planning (Gibbard 1990: 72-73). What interests me here is not so much Gibbard’s attempt to provide an evolutionary account of normative acceptance, but the general importance he attributes to the role of normative *language* in the development and functioning of moral practice.

The points I am about to make rest on what I take to be uncontroversial empirical claims about human nature and human society. Gibbard uses such points in the context of his broader evolutionary account of normative practice; I am interested in seeing how these facts about human nature and society might support the kind of approach to moral judgement I have been advocating. Human beings, it seems, are characterised by their capacity for language and by their sociality. And the human capacity for language would seem to figure centrally in the complex and sophisticated nature of human social life. It is via their use of language, says Gibbard, that human agents are able to engage in a complex system of shared evaluation of absent situations: to deliberate together about what to do and feel in such situations (1990: 72). If this is correct, and human evaluative practice has developed by way of and in conjunction with human language, then there should be a particularly tight connection between evaluative language and evaluative practice. In Gibbard’s terms, we might say that evaluative practice and motivation is ‘linguistically infused’ (1990: 57).

The idea here, as with the discussion of Strawson, is to see moral judgement in the context of its role in broader evaluative practice: a practice facilitated by the development of language. The starting point of this practice, we might say, is to find out what to do and how to feel in living our lives. There is thus a *presupposition of interest* that structures the practice, giving it its distinctive point. And if evaluative practice of this kind is to a large extent realised linguistically, then we would expect this presupposition of interest to be encoded in the rules governing the use of evaluative language and the performance of evaluative speech acts. And this is what

we find in the analysis of moral speech acts. The presupposition of interest is reflected in the preparatory condition for the speech act of moral evaluation: that the speaker is normatively engaged. It is also reflected in the illocutionary point or essential condition of moral speech acts: that the speaker's utterance counts as an undertaking to be guided by the content of the uttered evaluation.

Crucial to the human system of shared evaluation facilitated by language is the network of reactive attitudes and emotions. These attitudes share the same presupposition of interest that characterises moral speech acts. As we have seen, to be susceptible to such attitudes is just to be an interested participant in evaluative practice. Reactive attitudes are inherently evaluative in being characterised by central cases in which they are warranted responses to the violation of or conformity with practical evaluative requirements. These central cases thus involve evaluative, often moral, judgements. Because of the sophisticated, highly abstract, and shared nature of moral evaluations, it is by way of normative language and 'normative discussion', as Gibbard terms it, that human agents make such evaluations (1990: 73). The rules governing normative discussion therefore come to play a central role in what counts as normative or evaluative *practice*.

One reason moral judgements are language dependent is that what they are about – e.g. moral facts – are of such complexity that it would be *empirically* impossible to have moral thoughts without linguistic symbols (cf. Searle 1995: 64). Moral practice, by contrast, including the practice of making moral judgements, seems in part to be constituted by rules governing the use of moral language.<sup>3</sup> This has much to do with the inherently social and interpersonal nature of evaluative practice. Inherently social facts, suggests Searle, are by their very nature communicable and

---

<sup>3</sup> On Searle's conception of institutional facts, facts about evaluative and moral *practice* would count as institutional facts and thus language dependent (1995: chap. 3). He would also, I think, regard evaluative and deontic facts as essentially institutional and language dependent (e.g. 1995:70). While I am committed to viewing facts about moral practice as language dependent, this does not commit me to viewing moral facts as language dependent. In a similar way, I might view facts about scientific *practice* as essentially institutional and language dependent, without viewing scientific facts as language dependent. These considerations raise difficult ontological issues beyond the scope of the present work. But it does seem, as Searle admits, that there is not a sharp dividing line between institutional and non-institutional facts or between linguistic and pre-linguistic phenomena (1995: 71).

thus require a means of public communication such as a language (1995: 77). This inherently social or shared nature of moral practice is reflected in the reactive attitudes that form a central part of that practice. As we have seen, the reactive attitudes are inherently interpersonal: they essentially involve the assessment of the attitudes and behaviour of ourselves in relation to others and of others in relation to each other and ourselves.<sup>4</sup> The dependence of the reactive attitudes on the linguistically infused practice of evaluation suggests that such attitudes might be what have been termed ‘higher cognitive emotions’ (Griffiths 1997). Higher cognitive emotions such as guilt, shame, jealousy, and pride are contrasted with ‘basic emotions’ such as anger, fear, surprise, and disgust. While both types of emotion are said to be universal in humans, higher cognitive emotions are characterised by being ‘fundamentally social in a way that basic emotions are not’ (Evans 2003: 20).<sup>5</sup>

With the above considerations in mind we can again turn to the objection that the rules outlined in my account of moral speech acts are merely linguistic conventions that provide a trivial accommodation of the internalist requirement. As already mentioned, this argument stresses that it is always logically possible to envisage someone having a moral belief in the absence of normative expectation and in the absence of normative engagement. The above points about the connection between moral language and broader moral practice help to flesh out the response I made at the end of the last chapter. The rules governing moral speech acts reflect, in Strawson’s terms, a fundamental human commitment to the participant point of view. My contention is that moral practice includes a conception of making a moral judgement that involves holding agents responsible as well as asserting that some fact obtains. Strawson argues that susceptibility to the reactive attitudes and the

---

<sup>4</sup> Cf. Strawson: ‘[B]eing involved in inter-personal relationships as we normally understand them precisely is being exposed to the range of reactive attitudes and feelings...’ (1962: 54).

<sup>5</sup> Evans continues: ‘You can be afraid of, or disgusted by, inanimate objects and non-human animals, but love and guilt require other people for their existence (2003: 20-21). Gary Watson also stresses the communicative role of the reactive emotions, seeing the point of such attitudes as forms of moral address: ‘The reactive attitudes are incipient forms of communication, not in the sense that resentment et al are usually communicated; very often, in fact, they are not. Rather, the most appropriate and direct expression of resentment is to address the other with a complaint and a demand (Watson 1987: 127-128).



concomitant participation in evaluative practice is not something that human agents can choose to opt out of.<sup>6</sup> More fundamentally, even if it were possible to take up a wholly objective attitude towards oneself and others, one would no longer be taking up a recognisably moral point of view.<sup>7</sup> To give up on the linguistic rules that characterise moral practice is to give up on the norms governing the practice and the attitudes that are an integral part of it. To abstract oneself from these attitudes is to give up much of what it is to be a recognisably human agent.

### 3. Conative Motivation and Normative Motivation

Much of this work has been concerned with developing a plausible version of moral judgement internalism consistent with moral cognitivism. To some extent it has circumvented the issue, often taken to be central to cognitivist internalism, of whether moral beliefs can motivate independently of desire or play the dominant role in motivation. As we saw in chapter one, opposition to the notion of cognitive motivation is standardly based on adherence to the desire/belief model of motivation. Cognitivist internalists often, therefore, see their main task as one of undermining the desire/belief model (cf. Dancy 1993: 7-9). In contrast to such approaches, I have tried to develop a version of cognitivist internalism that does not rely for its *initial* plausibility on providing an alternative to the desire/belief model of motivation. It has become evident, however, that the picture of moral motivation suggested by the speech act account is potentially too complex to be accommodated by the desire/belief model. In light of this, I want to approach the issue of the desire/belief model by considering it in relation to a distinction between two kinds of motivation.

On the speech act model, an agent making a moral judgement is doing something more than merely describing some state of affairs. Moral judgements are active in the sense that an agent who makes a judgement commits herself to being guided by the judged requirement by holding herself and others to account with regard to that requirement. This captures a robust sense of endorsement absent in the case of

---

<sup>6</sup> See Strawson 1962: 54: 'A sustained objectivity of inter-personal attitude, and the human isolation which that would entail, does not seem to be something of which human being would be capable'.

<sup>7</sup> Cf. Strawson 1962: 59: '[T]o the extent to which an agent is seen [wholly in terms of the objective attitudes]...he is not, to that extent, seen as a morally responsible agent, as a term in moral relationships, as a member of a moral community.'

merely believing or stating something. As a consequence, the speech act approach is well placed to deal with the non-cognitivist anti-cognitivist argument discussed in chapter one. Central to that argument is the claim that purely cognitive states are incapable of motivating. Consider Dancy's presentation of a version of this argument:

I start with what I take to be a recognisable form of the Humean argument that moral judgements must be at least partly non-cognitive. If we add to a purely cognitive state a moral judgement, is the result a complex state that is 'sufficient for action'? This phrase 'sufficient for action' is itself one of the difficulties. But the question it is used to ask is not itself too obscure. It is whether starting from a state that is purely cognitive, and adding to that state a moral judgement, it is conceivable that an intentional action should then occur without further addition to, or other change in, the psychological state of the agent. Humeans take it that the answer to this question is yes...[but] that if the moral judgement is itself conceived as purely cognitive, the answer to our question would have... to be no. This is because they suppose that a purely cognitive state is incapable of motivating – is not and cannot be 'sufficient for action'- without the presence and help of an independent desire. (Dancy 2000: 82)

As I argued in chapter one, the appropriate conclusion to draw from this argument is that if moral judgements are cognitive, then they must involve something in addition to cognition. However, this does not mean that the cognitive *state* involved in moral judgement need be anything other than purely cognitive. The speech act approach defuses the anti-cognitivist potential of the argument by showing in what sense moral judgements go beyond cognition. And it does this without needing to question the Humean distinction between essentially motivating states like desire and essentially non-motivating states like belief.<sup>8</sup> However, I think we have reason to question the Humean account of motivation as it applies to intentional action. The Humean argument says that by simply adding a relevant desire state to a cognitive state we are in a position to explain motivation and intentional action. But on the model I have presented, the endorsing attitudes expressed in moral judgement seem to play a different role in an agent's motivational economy to the kind of motivation we have seen can fail as the result of depression, accidie and so on. This is reflected in the fact that while an agent who makes a moral judgement necessarily endorses that which is judged right, she is not necessarily motivated to do what she judges is right.

---

<sup>8</sup> As will become clear, I don't wish to question the Humean distinction between conative and cognitive *states*, but the notion that the sense of 'motivation' essential to conative states is the only sense of 'motivation' we make use of in explaining intentional action.

To understand how different kinds of motivation can fulfil different roles in the way an agent acts, consider the difference between the attitudes that make an agent a compassionate person and whatever it is that motivates a compassionate person in particular instances to, for example, help someone. Though what makes someone count as compassionate is, plausibly, a range of attitudes, such as sympathy and concern for others, that incline the agent to help those in need, the compassionate person may or may not be motivated by any particular instance of perceived need. When she is motivated, there will typically be features of the situation that motivate her concern: features she thinks warrant her compassion and her acting to help.<sup>9</sup> Typically, an agent will not choose when to feel compassion; but when she does feel compassion she will assess whether or not the compassion is a warranted response to the situation. If she judges her compassion warranted, then she will tend to be motivated to assist in some way the object of her concern. What is going on in such cases? I suggest there are actually two different senses of ‘motivation’ involved in such examples. Recall in chapter one I said that the term ‘motivation’ is ambiguous between motivating states and motivating considerations.<sup>10</sup> In chapter four, I called the former kind of motivation ‘non-normative motivation’ and the latter kind of motivation ‘normative motivation’. In the rest of this section I want to explore in more detail the differences between these two senses of ‘motivation’.

If an agent has a conative or non-cognitive state that inclines her to act in a certain way, we would standardly say that the agent is motivated to act in that way. Conative or non-cognitive states of this kind are said to have a desire-like world-to-mind direction of fit. If we think of desire as simply any state with which the world must fit, then we can say that to have a motivating state disposing one to act in a certain way is just to have a desire to act in that way.<sup>11</sup> For example, suppose that whenever John gets near the edge of a cliff or a tall building he is assailed by a

---

<sup>9</sup> Sometimes, of course, she may think her compassion unwarranted; but in general we would expect her compassion to track her beliefs about the kinds of things that warrant compassion.

<sup>10</sup> I also said that ‘motivation’ can sometimes refer to the state that conveys the consideration in light of which an agent is motivated. This sense of motivation only makes sense if the distinction between the other two senses is sound. So for now I will concentrate on these other two senses.

<sup>11</sup> Cf. Smith (1994: 117): ‘[I]f ‘desire’ is not a suitably broad category of mental state to encompass all of those states with the appropriate direction of fit, then the Humean may simply define the term ‘pro-attitude’ to mean ‘psychological state with which the world must fit’...’

strong desire to jump off. Merely having this desire means that John is, what I will term, *conatively motivated* to jump off the cliff.<sup>12</sup> To be conatively motivated to act in a certain way is just to be in a psychological state with a world-to-mind direction of fit that disposes one to act in that way. But there is another sense of the term ‘motivation’ according to which John is unlikely to be motivated to jump off the cliff. Suppose John in no way *endorses* his desire to jump off high places: he does not think the desire provides justifying grounds to jump off tall buildings and cliffs. Moreover, suppose his tendency to experience such desires actually motivates him to stay away from such places, and that it does this by being a consideration he takes to favour staying away from such places. To be motivated in this way is to be, what I term, *normatively motivated*.<sup>13</sup> An agent is normatively motivated to act in a certain way if she is motivated by a consideration she takes to favour or justify acting in that way.<sup>14</sup> I will mean by talk of justifying *considerations* or the *considerations* that normatively motivate an agent, the content of an agent’s belief - what is believed by the agent (Cf. Dancy 2000; 113) .<sup>15</sup> So if John thinks that avoiding rush hour counts in favour of leaving work early, then what he believes, ‘that he will avoid rush hour,’ is the consideration he takes to favour leaving early.

---

<sup>12</sup> This is a more precise term for what I previously called ‘non-normative motivation’.

<sup>13</sup> I borrow the expression ‘normative motivation’ from Gibbard (1990: 56). As a non-cognitivist, he obviously has a different account to me of the nature of this kind of motivation. Nonetheless he too uses it in a way that is intended to capture the intuitive distinction between motivation an agent endorses and motivation an agent does not endorse.

<sup>14</sup> Insofar as an agent acts for what she takes to be bad reasons she acts irrationally and lacks, it would seem, a degree of control over her actions. Practical irrationality plausibly involves an awareness of what one is doing without full control over what one is doing. John is aware of his desire to jump off the cliff and that he thinks it provides him with no justification for jumping off the cliff. Nevertheless, when he is unable to resist the desire he does jump off the cliff rather than merely fall: he is aware of what he is doing but is compelled in a way that is not consistent with fully autonomous intentional action. Such cases seem to represent a grey area between an agent’s being unconsciously caused to behave in a certain way and being motivated to act on the basis of her rational judgement. I take it that the central and defining cases of intentional action are those cases where an agent acts for a reason she takes to favour the action. It is the motivation involved in such cases that I am calling normative motivation. Velleman argues that when agents act for a reason, the reason need not be presented as a good thing to do, although it must be presented as justified (1992: 21). I would say that judging something to be a justifying reason for action is just to judge it a good reason for action. In this sense, at least, an agent who acts on such a reason would seem to be acting on something she takes to be good.

<sup>15</sup> If justifying considerations are what is believed by an agent then it seems they must be propositions. I share some of Dancy’s (2000: 114) misgivings about thinking of justifying considerations in this way, but I will adopt this position for the purposes of this chapter as a full discussion of these issues is

Consider again the Humean claim that only desires are capable of motivating agents. Given the above distinction, we should now ask whether the Humean is saying that only desires are capable of conatively motivating agents or whether he is saying that only desires are capable of normatively motivating agents, or whether, perhaps, he is making both these claims. If the Humean is arguing that only desires can conatively motivate agents, then I think we should agree with him, given our broad definition of desire. This means that someone is conatively motivated to act in a certain way if and only if she has a desire to act in that way. Following Dancy, we might say that to have such a desire is just to be in the state of being motivated (2000: 14). The claim that only desires can normatively motivate is, I think, far more contentious. It means that only facts about what an agent desires can count as justifying grounds capable of motivating her to act. On this model, the consideration that someone needs help, for example, cannot itself be a consideration I take to favour helping that person and which motivates me to help.

It is crucial to note that even if agents are only ever normatively motivated by their desires, this does not undermine our original distinction between normative and conative motivation. Consider the following example. Suppose I desire to drink a cup of coffee. This means that I am in a state that disposes me towards having a cup of coffee. That is, I am conatively motivated to have a cup of coffee. Let's say I have not decided whether or not I ought to have a cup of coffee, so I am not yet normatively motivated to have a coffee. After considering the matter, I decide that given I have only had one cup of coffee in the day my desire for coffee is reasonable grounds for having another cup. So the consideration that I desire a coffee succeeds in normatively motivating me to have another cup.<sup>16</sup>

Presumably, if I am normatively motivated by the consideration that I desire a cup of coffee (call this desire, desire (1)), then I form another desire to have a coffee (call

---

beyond the scope of the present work. This view of considerations enables me to make the distinction I am concerned to make between normative and conative motivation.

<sup>16</sup> If I had thought that two cups of coffee would keep me awake all night, I might have decided it best not to act on my desire to have another cup of coffee. In this case, my desire is not itself a reason against having a cup of coffee (in the way John's desire to jump was a reason to stay away from high places); the reason against having another coffee is that it would prevent me from sleeping.



this motivated desire, desire (2)). Now, the roles of these two desires in motivating me are obviously quite different. Following Schueler's use of a distinction he derives from Nagel, we can say that desire (1) is an 'unmotivated desire' and that desire (2) is a 'motivated desire' (See Nagel 1970: 29; Schueler 1995: 15-28).<sup>17</sup> Desire (2) is motivated in the sense that it is a desire I hold for a reason as a result of deliberating about what to do and making a decision on the basis of that deliberation.<sup>18</sup> That is, it is a desire I have because I take it that there is something that favours doing what the desire is a desire for. I take it that there are justifying considerations that favour satisfying desire (1) and this motivates me to pursue its satisfaction, resulting in desire (2). So desire (2) is explained in terms of my reason for holding it, namely, that I have a (reasonable) desire for a cup of coffee. Desire (1) is not formed on the basis of deliberation and is not held for any reason.<sup>19</sup> It is in this sense that it is unmotivated. Although desire (2) is a motivated desire, when I have that desire I am in the state of being motivated, which is to be conatively motivated. Conative motivation is common to both motivated and unmotivated desires. A motivated desire is the state that results when an agent is normatively motivated.<sup>20</sup>

---

<sup>17</sup> My account here is greatly influenced by Schueler's discussion of these issues (1995). Although he does not make a distinction between normative and conative motivation, the distinction is suggested by his account of the difference between motivated and unmotivated desires, and by his 'deliberative model' of action explanation (Schueler 1995: especially chaps. 1 and 6)

<sup>18</sup> Schueler notes that if an agent decides against a course of action, like having a coffee, then she typically will not have the motivated desires she would have were she to pursue that course of action (1995: 22-23). Such an agent will not, that is, form a desire to boil the kettle or a desire to get the coffee from the cupboard, and so on.

<sup>19</sup> There will obviously be a physiological explanation of my having this desire which will be one kind of reason for the desire, though not the kind of favouring reason I am concerned with here.

<sup>20</sup> Schueler (1995: 18) distinguishes between Nagel's claim about the difference between motivated and unmotivated desires and, what he terms, Nagel's 'entailment point': 'that *whatever* may be the motivation for someone's intentional pursuit of a goal, it becomes in virtue of that pursuit *ipso facto* appropriate to ascribe to him a desire for that goal (Nagel 1979: 29). This ascribed desire need not, it seems, be a genuine desire state; rather it is a kind of 'placeholder' or a way of talking about the fact that the agent is motivated. Schueler calls this the 'pro-attitude' sense of desire (1995: 34 – 35). The distinction between pro-attitudes and proper desires is central to Schueler's general thesis: that intentional action need not always involve the presence of a desire (proper). I do not consider this position in the text. It seems to me that one can produce a similar criticism of the desire/belief model by noting that a proper desire may or may not be part of what an agent deliberates about, even though some proper desires are entailed by intentional action because a desire is just the state of being motivated (cf. Dancy 1993: 7-9; 2000: 13).



The Humean claim that beliefs cannot motivate without desires and cannot be the primary source of motivation seems true in the case of conative motivation. To be conatively motivated is to be in a state with a world-to-mind direction of fit such as a desire. Beliefs cannot be internally motivating states in this sense because they have the wrong direction of fit (cf. Dancy 1993: 3). However, the kind of state needed for normative motivation to take place is a state that can convey the consideration in light of which an agent is motivated. And it would seem that a state with a mind-to-world direction of fit such as belief is ideally suited to fulfil this role. In the next section, I explore these considerations in relation to the desire/belief model of action explanation.

#### **4. Normative Motivation, Intentional Action, and the Desire/Belief Thesis**

It might be thought that the notion of normative motivation is mysterious and that it should be rejected because inconsistent with the standard desire/belief model of motivation and action explanation.<sup>21</sup> In this section, I argue that the concept of normative motivation is crucial to our notion of intentional action, and that to deny the possibility of normative motivation is to threaten the possibility that agents act for reasons. If we accept the basic distinction between the two senses of ‘motivation’, we need to determine whether the desire/belief thesis is primarily a thesis about conative motivation or about normative motivation.<sup>22</sup> I argue that desire functions in the desire/belief model conatively, as a motivating state, and not as a consideration in light of which the agent is motivated. However, I claim that insofar as the desire/belief thesis concerns the explanation of intentional action, the desire in question must be a motivated rather than an unmotivated desire. Given this, the real explanatory power of desire/belief accounts of intentional action derives from the considerations in light of which an agent is motivated to act and not from the psychological states an agent is in when she is motivated to act.

---

<sup>21</sup> Mackie, in his argument from queerness, suggests that the notion of evaluative considerations motivating agents is highly mysterious (1977: 38-42)

<sup>22</sup> I say ‘primarily’ because even if desire/belief thesis primarily concerns normative motivation, the fact that it involves an agent’s being motivated means that the agent will be in a motivating state and so be conatively motivated to perform the relevant act.

The desire/belief thesis is one way of understanding the Humean doctrine that all motivation has its source in desire. It is often put forward as a theory of intentional action explanation *and* as a theory of motivation. In general there is little or no distinction made between intentional action explanation and the explanation of motivation. Yet, *prima facie*, there is a distinction to be made because it is possible that an agent is conatively motivated to do something without actually doing it or even intending to do it. So it seems that intentional action cannot always be explained by whatever explains motivation. For example, given we know that John does not endorse his desire to jump off the cliff and that the desire is an unmotivated desire, our explanation of his being conatively motivated to jump of the cliff will presumably appeal to physiological considerations (he might have vertigo) or perhaps to unconscious psychological factors (perhaps he had a bad experience with heights as a child). It seems clear that the desire/belief model is not intended to account for this type of unmotivated desire: one does not need to explain an unmotivated desire to do something by invoking another *desire* plus some related belief that produces it. The desire/belief thesis is not a theory about the generation of unmotivated conative states.

Presumably, the kind of motivation the desire/belief model aims to account for is the kind of motivation closely tied to intentional action. Intentional action is action done for a reason; the desire/belief model aims to give an account of the motivation that occurs when an agent acts for a reason. Put another way, we might say that the desire/belief model aims to explain intentional action by citing the reasons that motivate an agent to act. However, the debate about the nature of intentional action explanation is complicated by there being two apparently quite different accounts of what we are asking when we ask why an agent acted as she did. Some theorists think that we answer this question by citing the psychological states the agent is in when she is motivated to do something intentionally. Other theorists argue that we answer the question by citing the considerations in light of which the agent is motivated to act. Smith, for example, appears to interpret the desire/belief model of action explanation as a model of the states an agent is in when intentional action occurs. The explaining reasons given by this model, says Smith, are '*psychologically real*'

(1994: 96). As such, these reasons are ‘*psychological states...that play a certain explanatory role in producing action*’ (Smith 1994: 96). By contrast, Dancy interprets talk of an agent’s reasons as talk of the ‘considerations in light of which he acted’ (2000: 2). On this interpretation, if the desire/belief model is a model of agents’ reasons, it says that the considerations that persuade agents to act in a certain way are always made up of desires and beliefs of the agents.

*Prima facie*, Dancy and Smith seem to be involved in different tasks. Smith is offering a theory of motivating states and Dancy is offering a theory of motivating considerations. Smith’s point appears to be that we explain intentional action by citing the conative and other states an agent is in when she acts intentionally. Take the above example of my having a cup of coffee. Let’s interpret it in terms of Smith’s claim that explaining reasons are psychological states present when an agent acts intentionally. This claim, says Smith, is captured by Davidson’s formulation of the desire/belief model of action explanation (Smith 1994: 92). Davidson gives the following formulation of the desire/belief model:

C1 R is a primary reason why an agent performed an action A under description d only if R consists of a pro attitude [desire] of the agent towards action with a certain property, and a belief of the agent that A, under description d, has that property. (Davidson 1980: 5)

Let’s say that the relevant description d of the action A in the coffee example is ‘having a cup of coffee’. So according to C1, my primary reason for having a cup of coffee must consist of a desire to have a cup of coffee plus a belief that to perform action A is just to have a cup of coffee.<sup>23</sup> Smith argues that this kind of desire and belief combination provides necessary and sufficient conditions for the existence of explaining reasons or reasons that explain the agent’s action (1994: 96). This means that positing such states must have a high degree of explanatory power: they must be able to explain why an agent acts as he does without recourse to further or different

---

<sup>23</sup> Schueler argues that it is unclear from the text of Davidson’s ‘Actions, Reasons, and Causes’ whether C1 plus C2, the claim that a primary reason for an action is its cause, are both necessary and sufficient for explanations of action in terms of an agent’s reasons or only necessary conditions (see Davidson 1980: 12; and Schueler 1995: 44 and note 1.) Smith argues that his version of C1 gives necessary and sufficient conditions for the existence of explaining reasons (1994: 96).

explanations. Yet it is not at all clear that positing such a combination of states provides a convincing explanation of intentional action.

Consider again the coffee example. If we take my unmotivated desire (1) to have a cup of coffee plus a belief, which I presumably have, about what constitutes having a cup of coffee, then according to Smith we have a combination of states that satisfies C1 and can thus explain my having a cup of coffee if I in fact do. But this seems wrong. To get from desire (1) plus the relevant belief to intentional action, I must decide to act on desire (1) and be motivated to do so.<sup>24</sup> This is what produces the motivated desire (2): the desire held for a reason or on the basis of deliberation. And it would seem to be a process that includes a decision of this sort and the related motivation that allows us to distinguish intentional action from an agent's merely being non-rationally caused by psychological states to behave in a certain way. The desire (1) to have a cup of coffee plus the relevant belief might, of course, cause me to have a cup of coffee even though I have decided not to have another cup. In this case, we might say that the reason I have a cup of coffee is the presence of a desire for a cup of coffee plus a relevant belief. But 'reason' in this sense is not the sense appropriate for intentional explanations. In the purely causal sense of reason, it might have been the case that the unmotivated desire to have a cup of coffee plus the relevant belief caused me to eat a phonebook and that they were the 'reason' I ate the phonebook. But this would presumably not mean that I intentionally ate the phonebook.

Korsgaard, discussing an example of Nagel, makes what I take to be the same point about the difference between intentional action and the non-rational causation of behaviour:

In Nagel's...example, a person has been conditioned so that whenever he wants a drink and believes the object before him is a pencil sharpener, he wants to put a coin in the pencil sharpener. Here the co-presence of belief and desire reliably lead to a certain action, but the action is a mad one. What is the difference between this person and one who, rationally, wants to put a coin in a soda machine when she wants a drink? One may be tempted to say

---

<sup>24</sup> This would seem to involve something more than Smith's view that the agent 'put the relevant desire and belief together' (1994: 92). Presumably I can understand how to satisfy a desire without being motivated to act on that desire.

that a soda machine, unlike a pencil sharpener, is the source of the drink, so that the right kind of conceptual connection between the desire and the belief obtains. But so far that is only to note a fact about the relationship between the belief and the desire themselves, and that says nothing about the rationality of the *person* who is influenced by them. If the belief and desire still operate on that person merely by having a certain causal efficacy when co-present, the rational action is only accidentally or externally different from the mad one... So neither the joint causal efficacy of the belief and the desire, nor the existence of an appropriate conceptual connection between them, nor the bare conjunction of these two facts, enables us to judge that a person acts rationally. For the person to act rationally, she must be motivated by her *recognition* of the appropriate conceptual connection between the belief and the desire. We may say that she *herself* must combine the belief and the desire in the right way. A person acts rationally, then, only when her action is the expression of her own mental activity, and not merely the result of the operation of beliefs and desires *in* her. (Korsgaard 1997: 221)

Using the notion of normative motivation, we might summarise Korsgaard's point in the following way. If there is such a thing as genuine rational or normative motivation, then it must be possible that agents are sometimes motivated by the consideration that something is a reason in favour of acting in a certain way.<sup>25</sup> That is, agents must be capable of being motivated by the normative content of their practical judgements.

Does the notion of normative motivation rely on a substantial metaphysical conception of freedom of will? If some form of compatibilism, such as Strawson's, can be made to work, then it will accommodate the distinction between normative deliberation (and motivation) and compelled behaviour of various kinds without relying on a substantial metaphysics. I take it that Strawson claims there are standards of normative competence internal to the participant point of view that allow us to distinguish between someone capable of normal interpersonal relationships, capable of holding themselves and others accountable, and someone who is incapacitated in such a way that they cannot act on the basis of deliberation (1962: 54-55). These latter individuals are those to whom it is appropriate to take up the objective point of view. However, if compatibilism fails and libertarianism proves untenable, this would not mean the vindication of a purely conative version of the desire/belief model. Rather, because the notion of intentional action, or acting on the basis of

---

<sup>25</sup> Korsgaard argues that this is even the case with the instrumental or means-end principle (1986: 13). Her point seems to be that in order to be *rationally* motivated by the consideration that x is a means to achieving a desired end y, an agent must be moved by the consideration that x being a means to y favours her doing x.



reasons, is tied up with ordinary notions of responsibility and deliberative agency, we would not have a model of intentional action at all. Normative motivation would be like the appearance of moral cognitivism to the moral non-cognitivist: something essentially illusory to be made sense of, ultimately, in other terms. My point is that insofar as we want to explain ordinary intentional action, and not explain it away, we will have to rely on a conception of normative motivation.

If intentional action is action motivated by favouring considerations and not merely behaviour produced by the presence of psychological states, then unmotivated desires will have little power to explain intentional action independently of motivated desires.<sup>26</sup> This means in the coffee example that positing desire (1) plus a relevant belief will not be sufficient to explain my intentionally having a cup of coffee.<sup>27</sup> One option for rescuing C1 might be to interpret the relevant desire as a motivated rather than an unmotivated desire. If we add this motivated desire to a relevant belief, this might give us a combination of *states* capable of explaining intentional action. A motivated desire is by definition a desire explained in terms of the agent's reason for holding it (Schueler 1995: 24-25). Given that intentional action is action done for a reason, motivated desires seem the right kind of psychological states to indicate the presence of intentional action.

However, this kind of move only serves to strengthen the primacy of explanations of intentional action in terms of the considerations in light of which agents act rather than in terms of psychological states. It would seem that acting for a reason, in Dancy's sense of an agent's reasons, is what distinguishes intentional from non-

---

<sup>26</sup> It would be extremely difficult, for example, to predict whether or not someone will have a cup of coffee based only on her having an unmotivated desire and a relevant belief about what action constitutes having a cup of coffee. Without knowing more about the agent and her situation, one could at best make a guess. The point is perhaps more obvious in the case of John's unmotivated desire to jump off the cliff. Recall that having this desire actually motivates John to stay away from exposed high places like cliffs. So basing a prediction that John will jump off a cliff on the fact that he has a desire to jump off a cliff and a relevant belief about how to do this would lead one to entirely the wrong conclusion.

<sup>27</sup> Neither, it would seem, is desire (1) a necessary condition of my intentionally having another cup of coffee. Suppose I do not have an unmotivated desire, such as desire (1), to have a cup of coffee because I do not like coffee very much. But say a friend is relying on me to stay awake so I can pick her up from the train station late at night, so I decide to have some coffee in order to stay awake. This produces in me a motivated desire to have a cup of coffee. I then go ahead and have a cup of coffee in the absence of any unmotivated desire of type (1) to have a cup of coffee.



intentional action. A revised version of C1 that posits a motivated desire plus a relevant belief is parasitic in terms of explanation on a prior notion of motivation by favouring considerations or normative motivation. It sets out the states involved in *being* motivated to act intentionally – the states involved in the conative motivational aspect of intentional action – but relies for its explanatory power on those states having been motivated by considerations favouring the action.

Dancy calls the view that agents' reasons for action are always psychological states of the agents 'psychologism' (2000: 14-15). He argues, rightly I think, that such states are rarely considerations in light of which agents act (Dancy 2000: 15). At times, however, I think Dancy overstates his position: 'If our motivating reasons are all 'what is believed' [rather than a belief of the agent], no reasons are psychological states of the agent...' (2000: 99). What Dancy is objecting to here is the view, exemplified by Smith's position, that intentional action is to be explained by citing the psychological states *an agent is in* when she is motivated to act intentionally. In the coffee example, this view says that we explain my having a cup of coffee by citing my desire (2) to have a cup of coffee and my belief that action A counts as having a cup of coffee. I do not think Smith thinks these explanatory states need necessarily be part of the content of the agent's deliberations (Pettit and Smith 1990). These are the states standing 'behind' action. Dancy is opposed to this view, but mistakenly thinks that if the proper form of intentional explanation does not involve the states behind action, then the considerations in light of which agents act are never psychological states of the agent. This seems false. For example, we have already seen that my desire (1) for a cup of coffee can be a consideration or reason in favour of my having a cup of coffee, and that John's tendency to desire to jump from high places can be a consideration or reason in favour of his staying away from high places. Psychological states, like many other things, can be part of what agents deliberate about.

A good deal of the confusion here would seem to stem from the notion of 'motivating reasons'. I have tended to use the expression 'explaining reasons' precisely because it wears its ambiguity on its sleeve: there are many kinds of reason

explanations that have nothing to do with intentional action (see Schueler 1995: 62-63). This forces us to be clear about how we are using 'explaining reasons' in the context of intentional explanation. By contrast, 'motivating reasons' gives the impression of being essentially about the reasons involved in intentional action. But it inherits the ambiguity of the term 'motivation'. This means we need to ask whether talk of motivating reasons is talk of conatively motivating reasons or of normatively motivating reasons. And this is just the difference between explanation in terms of the psychological states an agent is in when she is motivated and explanation in terms of the considerations in light of which she is motivated. And as we have seen, explanation of intentional action in terms of the psychological states an agent is in when she is motivated to act is parasitic on explanation in terms of the considerations that normatively motivate the agent to act. Motivation by considerations that favour action is what distinguishes the motivation involved in intentional action from behaviour non-rationally produced by psychological states.

Smith attempts to avoid the limitations of explanations that simply posit psychological states by construing the desire/belief model as a model of teleological explanation. Teleological explanations, he says, are 'explanations that explain by making what they explain intelligible in terms of the pursuit of a goal' (1994: 104). Now, teleological explanations do have a fair degree of explanatory power with regard to action explanation. But insofar as an agent's pursuit of a goal is intentional, the agent must be normatively motivated to pursue whatever the goal is and not merely non-rationally compelled by psychological states, or anything else, to behave in a certain way. If my goal is to have a cup of coffee, then I must have decided to have a cup of coffee and have a motivated desire to have a cup of coffee. These facts can begin to explain why, if I do, I have a cup of coffee in a way the mere fact that I have an unmotivated desire to have a coffee, plus a relevant belief, cannot.<sup>28</sup> Note that the fact that I have an end that I am normatively motivated to pursue can also explain why I perform actions that subserve that end in a way that

---

<sup>28</sup> The explanatory impotence of such desires is in part due to the fact that they do not, as they stand, represent the goals of the agent. To think otherwise would commit us to thinking that, for example, one of John's goals is to jump off cliffs, or that any transitory or trivial desire that might assail an agent reflects a goal of that agent. But in this case, we may wonder, as Korsgaard suggests, if there is anything left of the notion of agency (1997: 247).

unmotivated desires cannot. If my goal is to have a cup of coffee, then I will typically have motivated desires to get the coffee from the cupboard, boil the kettle, get out the milk, and so on. Motivated desires point to the considerations that have normatively motivated the agent to act: the agent's reasons. If, ultimately, the consideration that motivates the agent is an unmotivated desire, such as the desire for a coffee, this desire still functions as a consideration in light of which the agent is normatively motivated.

The Humean conception of the desire/belief model says that we can explain the motivation involved in intentional action and thus intentional action by laying out the constitution of the motivating states an agent is in when she acts intentionally. The simplest version of this model argues that any desire and related belief will potentially be able to explain intentional action. We saw that this cannot be correct as the presence of an unmotivated desire, and a related belief, is neither necessary nor sufficient to explain intentional action. A more sophisticated version of this model cites motivated desires plus relevant beliefs to explain intentional action. However, this version relies on and does not explain the kind of normative motivation that characterises the motivation involved in intentional action. Citing a motivated desire simply points us towards the considerations in light of which the agent is motivated, and these considerations do most of the explanatory work.<sup>29</sup>

Thinking about the problem of moral motivation in terms of whether desires or beliefs are the primary source of motivation obscures a more fundamental distinction between conative and normative motivation. Normative motivation is only possible

---

<sup>29</sup> Even citing a motivated desire plus a relevant belief will often be less helpful in explaining intentional action than one might think. This is because agents often have competing motivated desires. For example, Charles is motivated by considerations favouring going out and considerations favouring staying at home. Charles needs to weigh up competing considerations and make a decision about what to do. If we think about the kind of motivating state Charles is in once he decides what to do, it will be close to what Schueler terms an 'intention-generated motivated desire' (1995: 24). These are desires 'which exist in virtue of some decision or intention of [an agent] and [her] beliefs about how to carry it out...' (Schueler 1995: 24). Now, citing this kind of desire in conjunction with a relevant belief or beliefs will have a high degree of explanatory power with regard to intentional action. But, once again, the explanatory power of this kind of state is entirely dependent on its being the product of normative motivation, in this case, overriding normative motivation. The real force of intentional action explanation, then, comes from citing the considerations in light of which the agent decides to act, not the states an agent is in once she has decided to act.

if states with a mind-to-world direction of fit – cognitive states – convey the considerations that motivate. Now, if this form of motivation were unique to morality and unlike standard forms of motivation, then we would perhaps have reason to be sceptical about it. Dancy worries, for instance, that cognitive theories of motivation that treat moral or prudential motivation as a special case of non-Humean motivation existing alongside standard cases of Humean motivation are unacceptably hybrid (1993: 18-21).<sup>30</sup> But we have seen that our ordinary notion of intentional action depends on a conception of motivation akin to normative motivation.<sup>31</sup> If intentional action is possible then agents are sometimes motivated by favouring considerations, and moral motivation can be seen, not as anomalous, but merely as one form of normative motivation.

### **5. Internalism, Practical Force, and the Role of Conative States**

One way of viewing moral judgement internalism is as an attempt to secure the motivating force of morality. If moral judgements entail motivation, then it is impossible for agents to fail to be moved by what they judge to be moral considerations. And if moral judgements entail motivating states, it is thought these states can take part in desire/belief explanations of action: there can be moral action as a result of moral motivation. On this interpretation of moral judgement internalism, because agents are necessarily moved to do what they judge is morally required, there is a sense, it would seem, in which internalism secures the action-guiding function of morality as well as, or as part of, its motivating force. And this comes close to saying that moral judgement internalism secures the normative force of morality, if we think of moral considerations as practical norms. In this section, I argue against attempts to account for the practical force of morality by invoking the motivating force of conative states. If morality has practical normative force, it

---

<sup>30</sup> His concern is that such theories posit beliefs that sometimes do and sometimes do not need the contribution of a desire to motivate, and that beliefs 'are not carved up into two sorts, those that can motivate alone and those that need some help' (Dancy 1993: 21).

<sup>31</sup> The fundamental issue, then, is not whether beliefs or desires motivate, but the kind of motivation they are involved in and the contribution they make to that kind of motivation. Beliefs and desires can both be objects of normative deliberation and thus both can potentially normatively motivate agents. But because of the difference in their directions of fit, desires but not beliefs can conatively motivate, and beliefs but not desires can represent as true the considerations that favour action. Desires and beliefs, with their standard 'Humean' directions of fit, are thus perfectly suited to playing the roles required to motivate intentional action.

cannot be secured merely by positing non-normative conative states. This means that if moral judgement internalism is simply the thesis that moral judgements entail conative motivation, it will be unable to account for the practical force of such judgements.

Mele has dubbed the argument that moral judgement internalism is required to accommodate the practical or action-guiding nature of morality ‘the argument from toothlessness’: ‘It is alleged that to reject...internalism is to drive a wedge between moral judgements and intentional conduct, thus taking the essentially practical bite out of morality’ (1996: 737). We saw in chapter one that part of the appeal of non-cognitivism is that it appears to be able to account for the action-guiding and motivating aspects of morality. However, in this chapter I have argued that even the non-cognitivist will have to account for the difference, or apparent difference, between those desires an agent endorses and those desires she does not endorse (see 5.2). Moral judgements express an agent’s endorsement or commendation of an option, and the mere presence of a motivating state does not secure this form of endorsement.<sup>32</sup>

Non-cognitivism aims to accommodate the cognitive appearance of moral judgements and, presumably, the appearance that moral judgements have normative force capable of guiding agents in their actions. But non-cognitivism denies the reality of moral facts and properties and that moral judgements aim to describe such facts and properties. So non-cognitivism does not explain as much as explain away the action-guiding force of moral judgements. If agents are not really moved by the normative content of moral judgements, then what is left to explain is the fact that moral judgements reliably conatively motivate agents to act. Thus, when non-cognitivists talk about explaining the actual motivating force of moral judgements,

---

<sup>32</sup> A sophisticated version of non-cognitivism may be able to account for the apparent distinction between merely desiring to perform an act and endorsing the desire to perform an act (e.g. Gibbard 1990: chap. 4). However, following Brink, I think that we should treat moral cognitivism as the default account of moral judgements (Brink:1989: 25-29). Non-cognitivists argue that the inability of cognitivism to account for moral judgement internalism is one of the principal reasons for rejecting cognitivism about moral judgements. However, we saw in chap 1 that there is no quick way of dismissing the possibility of combining moral cognitivism and moral judgement internalism.



they must have in mind the conative force of such judgements. This is presumably why they argue that moral judgements are constituted by conative states. And, as I said above, if moral judgements entail conative states, it is thought that these states can take part in desire/belief explanations of intentional action. I have argued in this chapter, however, that the mere presence of an unmotivated desire is not capable of playing this kind role in the explanation of intentional action.

I want to illustrate this point again by looking at a response Mele says is available to the externalist confronted with the argument from toothlessness. He asks us to imagine a planet, planet X, on which the majority of Xians have ‘a long-term generic desire’ to do whatever is morally required of them. These beings, says Mele, ‘seemingly possess a practical, action-guiding morality. Except in rare cases, they have motivation to do what they apparently take themselves to be morally required to do – motivation with roots in the long-term desire (1996: 733-734, 737-738). Mele’s argument is that Xians are reliably motivated to do what is morally required even if moral judgement internalism is false: for Xians morality seems to have teeth. The problem with this argument is that the long-term desire in question is not capable, as it stands, of playing the kind of action-guiding role Mele ascribes to it. Mele says that the long-term desire ‘is a feature of the psychological constitution of Xians’ (1996: 733). It is a desire, that is, that the Xians have naturally; it is not something they choose. So the Xian long-term moral desire is clearly an unmotivated desire: a desire not held for a reason or as a result of deliberation. But unmotivated desires cannot normatively motivate unless agents take them as considerations favouring action. If the Xians act intentionally, they must take their generic moral desire as a justifying ground for doing what is morally required. If they are reliably normatively motivated by this desire, what accounts for this must be the content of their, apparently uniform, *judgement* that such a desire warrants being satisfied. Without a judgement of this kind, the unmotivated desire cannot do the work Mele wants it to do.

To see this, imagine a world in which all agents develop at a young age an unmotivated desire to do what is morally required but where the desire to do what is



morally required is uniformly viewed as trivial and unjustified. Perhaps such a desire is seen as dangerous, in the way we might view a widespread desire for unlimited power as dangerous. In any case, on this world when an agent judges that something is morally required, she will, given her standing desire, be conatively motivated to do what she judges is required. However, insofar as she is aware of the desire, she will judge it to be unwarranted and something to be overcome or suppressed.<sup>33</sup> So a standing desire to be moral is obviously not sufficient to account for the practical ‘toothiness’ of morality.

Mele’s Xian example shows why Xians are conatively motivated to do what they judge is morally required, but it does not accommodate the action-guiding or normative force involved in normative motivation. It points to a standing desire to do what is right that the Xians may or may not endorse. The model of moral judgement I have proposed aims to account not only for the cognitive nature of such judgements and for internalism about such judgements but to allow for the role of normative motivation or genuine action-guidingness in moral motivation. This means leaving room for the idea that normative force is something other than the production of behaviour by psychological states, and for the idea that normative motivation has a central role to play in the explanation of intentional action.<sup>34</sup>

---

<sup>33</sup> On the approach to moral judgement I have advocated, the agents on this world would not be making genuine moral judgements, as they would not be holding themselves and others to account. In this example, I am supposing externalism to be true and looking at the implications of its truth for the normative authority of morality. Externalists tend to identify judgements as *moral* judgements on the basis of such judgements being about certain things, such as human well being or other-regarding considerations, rather than on the agent’s attitudes towards the content of the judgement (such as the attitude of holding herself and others to account with regard to the content of the judgement) (see 2.5 for a discussion of the difference between content-based and functional views of moral judgement). On the externalist model, it is always possible to lack interest in or question the justification of particular considerations construed as moral. The problem for externalists is that an independent desire to be moral would, it seems, need to have some normative force if it were to play a role in intentional explanation. Another interpretation of the example is that the agents on the world reject the influence of the *unmotivated* desire to be moral because it distracts them from what they think should properly motivate agents to do what is right, namely, facts about moral rightness. In this case, although they view the unmotivated *desire* to be moral as trivial and unjustified, they might view morality itself as of supreme importance. Even on this interpretation, then, the unmotivated desire is not what explains the toothiness of morality.

<sup>34</sup> I say ‘leaving room’ because I am not here attempting the considerable task of saying what normativity is – although I am saying what I think it cannot be – or defending realism about normative facts and normative motivation. I am trying to point to some of the implications of denying the possibility of normative motivation in relation to the possibility of intentional action.

Korsgaard notes that it 'is common among empiricists to equate the question whether pure reason can be practical with the question whether we are ever motivated by belief alone', where motivation is seen purely as a causal relation between psychological states and actions (1997: 220). Implicit in the empiricist view is the reduction of normative motivation to conative motivation. We can understand this a little better if we reflect on Mele's comment that 'to reject...internalism is to drive a wedge between moral judgements and intentional conduct...' This suggests that moral judgement internalism is in part an attempt to secure the connection between moral judgements and motivating force by securing the connection between moral judgements and intentional action. The formulation of internalism Mele sets up to criticise demonstrates the empiricist point of view as applied to this problem. Here, I pick up on a number of points discussed at the end of chapter four with regard to moral belief internalism. Recall Mele's formulation of moral belief internalism:

Necessarily, any belief that one is (oneself) morally required to A constitutes motivation to A. (Mele 1996: 731)

What Mele means here is that moral beliefs constitute motivating *states*: that according to moral belief internalism moral beliefs must somehow have the direction of fit normally associated with desire-like conative states (1996: 740).<sup>35</sup> The problem with this kind of view should now be familiar. What it gives us is a state which both represents something as morally required and conatively motivates the agent who has the belief to pursue what is represented as required: 'a noncompound truth-seeking motivation constituting state' or 'besire' (Mele 1996: 744). Whatever we think of the possibility of such a state, it simply cannot fulfil the role it is meant to with regard to intentional action motivated by moral considerations. If an agent notes he has a besire that  $\Phi$ -ing is morally required then he can perhaps take the fact that he is in a conative state that inclines him toward  $\Phi$ -ing as justified by the consideration, represented by the same state, that  $\Phi$ -ing is morally required. This will typically mean he will be normatively motivated to  $\Phi$  and thus have a separate motivated desire to  $\Phi$ . That is, he will be motivated by the consideration that  $\Phi$ -ing

---

<sup>35</sup> Mele notes that internalists 'who equate motivation with desire are committed to the idea that some beliefs are identical with desires or, at least, encompass desires (1996: 730). He also notes that such states have been termed 'besires'. For discussions of the notion of besire see: Lewis 1988; Price 1989.

is morally required. The conative aspect of the besire adds nothing to the justification or to the normative motivating potential of the moral consideration represented by the cognitive aspect of the besire. Consequently, the desire-like direction of fit of the besire will play no role in the intentional explanation of an agent's doing  $\Phi$  in light of the consideration that it is morally required. This makes the notion of a besire unhelpful in attempting to secure the connection between moral judgements and intentional action.

In order to address the argument from toothlessness, the connection between moral judgements and intentional action needs to be accounted for. But this cannot be done merely by pointing to a conceptual connection between cognitive and conative states (the Xian example) or by making cognitive states partly conative (constitutive moral belief internalism). Both these examples suggest that normative motivation has a central role to play in any account of intentional action motivated by moral considerations. Even if moral judgements entail conative motivation, this does not, by itself, tell us how such judgements motivate intentional action.

## **6. Normative Content and Internal Reasons**

I said above that an adequate theory of moral judgement must account for the different ways motivation figures in the motivational economy of moral agents. The pressure towards internalism in part stems from a recognition that we have to explain the connection between the practical or normative force of moral judgements and the motivation involved in intentional action. Now, the obvious place to look for the normative or justifying force of moral judgements is in the normative content of those judgements. When an agent makes a moral judgement she asserts a normative proposition of the general form 'S's  $\Phi$ -ing is morally required' or 'S morally ought to  $\Phi$ .' This suggests that in order to explain the connection between normative force of moral judgements and intentional action, we need to explain the connection between the normative content of moral judgements and intentional action.

One might think that the failure of Mele's Xian externalism and of constitutive belief internalism to account for the *practical* normative force of moral judgements stems

from their failure to include motivation in the content of moral judgements. Whatever normative force Xian moral judgements may have, such judgements rely on a conative state entirely independent of them for their motivating power. While constitutive belief internalism posits moral beliefs that constitute motivating states, motivation is no part of the propositional content of such states. In chapter one, I argued against interpreting moral judgement internalism as a form of content internalism. However, it might be possible to defend content internalism as a separate thesis. If moral judgement internalism accommodates the endorsing or commendatory nature of moral judgements, perhaps content internalism as a distinct thesis can accommodate the practical normative force of moral judgements.<sup>36</sup> Below, I argue against this view: including a conative state in the content of a moral judgement can no more account for the practical normative force of such a judgement than adding an external conative state to a moral judgement or making moral judgements themselves partly conative.

Content internalism about motivation is intended to bridge the gap between, what Velleman terms, 'the story of motivation and the story of rational guidance' (1992: 3). That there is a gap is suggested by a common way of formulating the distinction between explaining and justifying reasons. This is the model advocated by Smith, according to which explaining reasons are viewed as psychological states and justifying reasons as normative propositions. The two types of reasons, says Smith, 'are of quite different *categories* (Smith 1994: 96).<sup>37</sup> We have already seen that for Smith explaining reasons are psychological states present when an agent acts intentionally. I argued that this view is unable to provide an adequate explanation of intentional action and is dependent for the explanatory force it does have on a prior notion of motivation by favouring considerations or normative motivation. Dancy objects to Smith's interpretation of the distinction between explaining and justifying

---

<sup>36</sup> I said in chap 1 that content internalism contradicts the influential doctrine that moral requirements are categorical imperatives that apply to agents regardless of their contingent desires. It is open to the content internalist to argue that we need to accept a revisionary account of moral requirements and moral content if we are to explain the connection between practical normative force and intentional action.

<sup>37</sup> Cf. Velleman 1992: 4: 'The desire and belief cited in [the] story [of motivation] are conceived as propositional attitudes...a reason for acting [in the story of rational guidance] is a proposition whose truth would...in some...sense justify an action.'

reasons because he claims it threatens to make it impossible for the reasons why an agent acts to be among the reasons in favour of action (2000: 104-105). He argues that an adequate theory of the relation between explaining and justifying reasons needs to explain how justifying reasons can contribute to the explanation of action (Dancy 2000: 103).<sup>38</sup>

The view that there must be an intelligible connection between justifying and motivating considerations is perhaps most associated with Williams' defence of an internal conception of justifying reasons (1981; 1995). Williams makes the same basic point as Dancy: that good reasons for action must be capable of motivating agents and of playing a role in the intentional explanation of action:

It must be a mistake simply to separate explanatory and normative reasons. If it is true that A has a reason to  $\Phi$ , then it must be possible that he should  $\Phi$  for that reason; and if he does act for that reason, then that reason will be the explanation of his acting. So the claim that he has a reason to  $\Phi$  – that is, the normative statement 'He has a reason to  $\Phi$ ' – introduces the possibility of that reason being an explanation; namely, if the agent accepts that claim (more precisely, if he accepts that he has more reason to  $\Phi$  than to do anything else). (Williams 1995: 39)

In order for justifying reasons to be explanatory in this sense, Williams argues they must be of following form: A has a reason to  $\Phi$  only if 'A could reach the conclusion that he should  $\Phi$  (or a conclusion to  $\Phi$ ) by a sound deliberative route from the motivations that he has in his actual motivational set [the agent's 'S'] – that is, the set of his desires, evaluations, attitudes, projects, and so on' (1995: 35).<sup>39</sup> By a 'sound deliberative route' Williams means a route that excludes all mistakes of fact and logic in reasoning about what to do. The external reasons theorist claims that an agent can have a good reason to  $\Phi$  even if he has no motivation in his motivational set that could lead him via a sound deliberative route to  $\Phi$ .

---

<sup>38</sup> My aim here is to link this issue to the problem about the connection between normative moral content and intentional action. I am assuming, therefore, that moral judgements have, or purport to have, normative or justifying content and that moral propositions are a type of practical normative proposition. Moral reasons, then, are, or purport to be, reasons in favour of action (see 2.5).

<sup>39</sup> Williams only concerns himself with the view that his internalist constraint on reasons is a necessary condition for an agent's having a good reason to act, although he thinks it also provides a sufficient condition (1989: 35-36).



Williams' central claim is that there are only internal reasons for action and that this places a motivational constraint on what can count as a good reason for action. His argument has generated a vast amount of literature on the possibility of external reasons and on the nature of justifying reasons.<sup>40</sup> Here, however, I want to focus on what he argues is one of the '*fundamental motivations* of the internalist account': the interrelation of explaining and justifying reasons (Williams 1995: 38. My emphasis.) My basic argument against Williams is that his account of internal reasons fails to explain how agents are motivated by justifying reasons in a way consistent with the desire/belief thesis.<sup>41</sup> I am assuming therefore that part of Williams' aim is to accommodate the desire/belief thesis and the associated notion of explaining reasons by including desire in the content of normative judgements. Williams does not mention the desire/belief thesis by name, but it is clear that he thinks that for an agent to be motivated to perform an act, the agent *must be in* a motivating (desire-like) state: 'When the reason is an explanation of his action, then of course it will be, in some form, in his S [his *actual* motivational set], because certainly – and nobody denies this – what he actually does has to be explained by his S (Williams 1995: 39).<sup>42</sup>

In what follows I assume Williams' motivational constraint on justifying reasons and, in the light of the constraint, consider three examples in which an agent, A, judges of herself that she has a good reason to  $\Phi$ . The first case is one in which A's judgement that she has a good reason to  $\Phi$  is true. On Williams' view, part of the truth conditional content of A's judgement must be that were she to make no factual or logical errors she would be motivated to  $\Phi$ . As her judgement is true, this condition is met (i.e. it is true that she would be motivated in such circumstances). In addition, A is occurrently motivated to  $\Phi$ . In the second case, A's judgement that

---

<sup>40</sup> For example: Korsgaard 1986; McDowell 1995; Parfit 1997.

<sup>41</sup> The view that Williams' account of internal reasons fails to bridge the gap between explaining and justifying reasons is suggested by Schueler (1995: 68-77). He stresses the importance of distinguishing between 'the facts or alleged facts referred to or described in the content of my reason judgement from the facts about me that are supposed to explain my action (Schueler 1995: 73-74).

<sup>42</sup> Williams' discussion of the gin and tonic example also suggests that he views action explanation along the lines of the desire/belief model: 'The agent believes this stuff is gin, when it is in fact petrol. He wants a gin and tonic...if he does drink [this stuff], we not only have an explanation of his doing so (a reason why he did it), but we have such an explanation which is of the reason-for-action form' (1981: 102).



she has a good reason to do  $\Phi$  is again true, and part of what A judges is that were she to make no factual or logical mistakes she would be motivated to do  $\Phi$ . But in this case A has no occurrent motivation to  $\Phi$ . In the third case, A judges that she has a good reason to  $\Phi$  and, again assuming Williams' constraint, part of what she judges is that were she to make no factual or logical mistakes she would be motivated to  $\Phi$ . However, in this case A's judgement is false because it is not true that she would be motivated to  $\Phi$  in the absence of mistakes of fact and logic.

Williams says that a constraint of adequacy on any account of how justifying reasons explain motivation is that the 'difference between false and true beliefs on the agent's part cannot alter the *form* of the explanation which will be appropriate to his action (1981: 102). Suppose that Jane believes the truck is about to hit her and this is why she moves out of the way. Williams' point is that the explanation of Jane's moving out of the way must be the same whether or not it is true that the truck is about to hit her. That is, we cannot say that when it is true that the truck is about to hit Jane the reason she moves out of the way is the *fact* that the truck is about to hit her, but when it is false that the truck is about to hit her the reason she moves out of the way is her *belief* that the truck is about to hit her.<sup>43</sup> Borrowing from our discussion of Smith's reliability thesis in chapter two, we might add another constraint on the adequacy of an account of how justifying reasons explain motivation: it must be consistent with the fact that agents' motivations reliably track changes in their judgements about what they have good reason to do.

Williams makes justifying reasons relative to an agent's motivational set so that aspects of that set can take part in the explanation of the agent's intentional actions. Consider this in light of our first example. A judges correctly that she has good reason to  $\Phi$  and so judges correctly that she would be motivated in suitable

---

<sup>43</sup> Williams would, I think, argue that the form-of-explanation constraint supports reason explanations that appeal to whatever is constant in this kind of scenario, namely the agent's psychological states. But we have already seen that action explanation solely in terms of the states an agent is in when she acts intentionally is inadequate to account for intentional action. I would say that the agent's reason, the reason that motivates the agent, in the truck example is the consideration, which she takes to be true, that she is about to be hit by the truck. This is Jane's reason both in the case where her belief is true and in the case where it is false. See Dancy (2000: chaps 5 and 6) for a detailed discussion of the problems raised by Williams' constraint. Also cf. Schueler 1995: 76.

conditions (no mistakes of fact or logic) to  $\Phi$ . A is as a matter of fact occurrently motivated to  $\Phi$ . Let's suppose that this is because she is in the relevant conditions and that her occurrent desire to  $\Phi$  is the desire referred to in the content of her reason judgement. So it seems that this desire can take part in a desire/belief explanation of her  $\Phi$ -ing if she does so: the desire is part of why she has a good reason and is part of the explanation for her acting for that reason.

In the second example, A's judgement that she has good reason to  $\Phi$  is true. This means it is true that she would be motivated to  $\Phi$  in suitable conditions. But A is not in the relevant conditions and has no occurrent motivation to  $\Phi$ . Now, for a desire conatively to motivate an agent, the agent must be *in* the desire state. So the desire A would have under suitable conditions is, in this instance, unable to play the conatively motivating role required of desire in the desire/belief thesis.<sup>44</sup> We might simply concede that agents can only be motivated by their reason judgements under suitable conditions. But it seems implausible to suggest that agents are only motivated by their normative judgements in cases where they make no mistakes of fact or logic.

In order to conform to the desire/belief thesis, it seems we must posit a motivating state other than the one mentioned in the content of the judgement to play the role of the motivating desire. But this means that the desire included in the content of the reason judgement in order to bridge the gap between justifying and explaining reasons loses its point: it no longer plays any role in the explanation of actions done for a good reason.<sup>45</sup> Moreover, if we opt to explain the agent's motivation by appeal to a desire that is not part of the content of her normative judgement, then our explanation violates the constraint about consistency of form: when the agent is in

---

<sup>44</sup> Recall that the desire/belief thesis is about the states an agent is *in* when she acts intentionally.

<sup>45</sup> This is basically Schueler's argument against Williams' doctrine of internal reasons: 'The original problem [that the internal reasons account was supposed to solve] was supposed to be that whatever facts give me a justifying reason to do something, these facts, just by themselves, won't explain my acting as I have a reason to act until we add some subjective elements, in particular, *both* my awareness of the facts in question *and* the associated desire of mine. This is not changed by the internal-reason theorist claiming that the facts that give me a justifying reason to act always include desires of mine...since a belief [that I have a reason and thus a desire], like any belief, could be false. I might not have the desire I believe I have (1995: 73).

the relevant conditions, her motivation is explained using the desire in the content of her normative judgement, but when she is not in these conditions her motivation is explained using another desire that is not part of her normative judgement.

In the third example, A's judgement that she has good reason to  $\Phi$  is false, and it is false because her belief that she would be motivated in suitable conditions to  $\Phi$  is false. If only occurrent desires can motivate, then it is obvious we cannot appeal to the desire in the content of a false judgement to explain why, if she is, A is motivated to  $\Phi$ . So, again, the desire that is internal to the reason judgement cannot fulfil its task of bridging the gap between justifying reasons and explaining reasons, and so it loses its point. And we have to look around for yet another desire and another form of explanation to account for the motivation, thus violating the consistency of form requirement.

It is surely a failure of any account of how justifying reasons explain motivation if agents can only be motivated by true reason judgements and only under ideal conditions. The reliability thesis tells us that agents' motivations reliably track their judgements about what they have good reason to do. If an agent judges that she has a good reason to  $\Phi$ , then, whether or not her judgement is true and whether or not she is ideally placed to make the judgement, we expect her, in the absence of tiredness, accidie and so on, to be motivated to  $\Phi$ . And if she changes her mind and judges that she has a good reason to do  $\Psi$  instead of  $\Phi$ , we expect her motivation to track her change of judgement, again independently of the truth or falsity of her judgement and independently of whether she is ideally placed to make the judgement. If only occurrent desire states are capable of motivating agents and of playing the role of desire in the desire/belief model, then including a desire in the truth conditional content of reason judgements will not succeed in accounting for the ability of justifying reasons to motivate agents and to play a role in the explanation of action.

Consider again Williams' requirement that if an agent has a reason to act, then it must be possible for the agent to act for that reason, and if he does act for that reason, then for that reason to be an explanation of his acting. We can meet this requirement using the notion of normative motivation. Suppose that Charles has a good reason to donate money to the disaster relief fund: the donation will contribute towards relieving the suffering of those affected by the disaster. If this is a good reason for Charles, then it must be possible for him to act for this reason. Let's say that Charles does donate to the disaster relief fund, and that the reason or consideration that motivates him to donate is that his donation will help to relieve the suffering of disaster victims. We explain Charles' intentional act of donation by citing his reason for donating: namely, to relieve the suffering of disaster victims. Thus, the reason in favour of donating is both the reason that motivates Charles to donate and the reason we can cite to explain his act of donating. Using the notion of normative motivation, we can explain how a reason that justifies an action both motivates and explains an agent's intentional performance of that action.

The above arguments suggest that the connection between justification and intentional action that internalism sees itself as establishing cannot be secured merely by making a conative state in some sense internal to a moral or normative judgement. And this is what we should expect if the distinction made between normative and conative motivation is correct. The connection between justification and the motivation central to intentional action is best understood in terms of the concept of normative motivation. The concept of normative motivation bridges the gap between justification and motivation precisely because it is the concept of being *motivated* by *normative* considerations.

## 7. General Internalist Disquiet

Certain versions of cognitivist internalism adopt what in chapter one I called a cognitivist theory of motivation. I used this description in relation to Nagel's account of moral motivation, which I take to involve the notion that agents are, or can be, normatively motivated by the justifying content of their moral judgements (Nagel 1970: 65). The notion of cognitivist motivation is somewhat misleading, however, as it is used to describe theories, such as Dancy's 'pure cognitivism', that hold that belief *states* rather than desire *states* are what motivate intentional action (see Dancy 1993).<sup>46</sup> It is clearer to call the position I advocate, and which I take Nagel to hold, a normative theory of motivation.

Normative motivation is characterised, as we have seen, by an *agent's* being motivated by considerations she takes to be justifying. When the agent's normative judgement is true and she acts on it, the considerations that justify her action can be said to motivate and to explain her action. The structure of this type of normative motivation and its relation to intentional action is captured in the following passage from Velleman:

The story of rational guidance tells how an agent acts for a reason. According to this story, a reason for acting is a proposition whose truth would reflect well on, count in favour of, recommend, or in some other sense justify action. A reason for performing the action exists so long as a proposition justifying the action is true. But an agent cannot act for this reason unless he has mental access to it – unless he believes the proposition or at least grasps it in some related fashion. And even if he has appropriately grasped the reason, and is therefore in a position to act for it, he doesn't ultimately act for the reason unless his grasp of it results in his being influenced or guided by its justifying force. An agent acts for a reason, then, when the action-justifying character of a proposition prompts his action via his grasp of that proposition. (Velleman 1992: 4)

While I substantially agree with this story, it misses out something crucial about practical judgements such as moral judgements: it misses out the endorsing aspect of moral judgements or the sense in which making such judgements goes beyond cognition. Those who accept this model of normative motivation appear to view normative judgements as purely descriptive. They also generally accept that agents

---

<sup>46</sup> Dancy repudiates this form of, what he terms, psychologism in *Practical Reality* (2000) and adopts the position that states of affairs or facts rather than psychological states are what motivate intentional action.

sometimes fail to be motivated by their normative or moral judgements. In the case of moral judgements this raises a problem. It makes it difficult to distinguish the extreme amoralist, who has none of the attitudes we would standardly associate with a normatively competent agent, from the normal moral agent who suffers a failure of motivation.

On the descriptivist view, the normal moral agent's judgement and the extreme amoralist's judgement are both purely descriptive: they both function merely to express the belief that something is morally required. On the basis of this purely cognitive criterion, it seems that both types of agents can make genuine moral judgements, the only difference being that the normal moral agent is reliably motivated by her moral judgements whereas the extreme amoralist is never motivated by her moral judgements. However, we presumably want to say that the normal moral agent in some sense endorses, or is *connected* to, her moral judgements even when she fails to be motivated by them, in a way that is clearly not true of the extreme amoralist. Darwall makes a similar point about moral judgements and normative judgements more generally:

If something's being a reason is simply a non-natural property of it of which we take notice in judging the consideration to be a reason, then the desire to act for reasons is in no sense integral to the self... So understood the desire to act for reasons is not in itself intelligible. We cannot see it as essential to us. *But...our disposition to look for reasons and to act on those we find compelling cannot be so easily separated from us.* (Darwall 1983: 58. My emphasis)

The problem with a purely descriptivist theory of moral judgement is that it leaves it unclear in what sense agents are connected to the reasons for which they act. And this is surely grist to the externalist mill.

Here I think we find a general source of internalist disquiet: the sense that if we don't guarantee the connection between morality and motivation, *or something internal to agents*, then there will be agents beyond the reach of moral norms. This is a worry I take to motivate both judgement and content versions of internalism. We find a particularly vivid expression of this disquiet in Falk's seminal defence of internalism about moral requirements and motivation. The problem with externalism, argues



Falk, is that it contradicts our ordinary thinking about the connection between moral requirements and motivation: 'For it is a fact about ordinary moral thinking that morality needs...no sanction whatever; that somehow the very fact of a duty entails all the motive required for doing the act' (Falk 1948: 121). According to Falk, moral requirements entail overriding justifying reasons for action, and this amounts to their entailing overriding motivation to act. In order to secure the overriding motivating force of moral requirements, Falk reduces the kind of, presumably normative, necessity expressed by such requirements to causal necessity. This is the only way, he argues, of securing the intuition that the existence of such a requirement for an agent 'is inseparable from the existence of some real check on his freedom to act otherwise' (Falk 1948: 122).

Falk suggests a meaning of 'ought' or 'duty' which he calls the 'purely formal motivation sense' (1948: 128). In this sense, to say that an agent morally ought to do an act, is to say that she is disposed to have a belief about doing the act which causally compels her to do the act. This is the one use of 'ought', says Falk, 'which could explain the habit of connecting it necessarily with motivation' (1948: 130). According to Falk, then, agents have a duty to do those things they cannot help but do because causally compelled to do by certain belief states. What is striking about this view is Falk's willingness to dispense with anything resembling a *normative* requirement in order to secure the connection between 'moral obligation' and overriding motivation. The corollary of Falk's view is that agents can never fail to do what they have a moral obligation to do because agents only have an obligation to do those things they will be compelled to do.

Falk's paper is important not so much because of the thesis he expounds but because it is here that he coins the terms 'internalist' and 'externalist' to refer to two opposing positions in ethics. However, Falk seems to have a far more literal view of what these terms mean than many of the theorists who have used them since:

[Externalists] presuppose...that when someone 'ought or 'has a duty' he is subject to some manner of demand, made on him without regard to his desires, and they imply that this demand issues essentially *from outside the agent*: that whether made by a deity or society, or the 'situation'...it has an objective existence of its own depending in no way on anything peculiar to the agent's psychological constitution. Now, the view that morality needs some sanction is a traditional associate of all views of this kind and indeed their natural corollary. If 'I ought' means 'I am from outside myself demanded to do some act'...*there will then be no necessary connection for anyone between having a duty and being under any manner of real compulsion to do the act*. For no one really need do any act merely because it is demanded of him...but only if, in addition, he finds within himself a motive sufficient for satisfying the demand. (Falk 1948: 125-126. My emphasis)

So for Falk, an external theory of moral obligation is one that views moral obligations as imposed on the agent from outside, whereas an internal theory of moral obligation views such obligations as inside the agent because part of the agent's psychology.

It is clear that Falk dismisses the possibility of forms of necessity other than causal necessity and that the only sense of 'motivation' he recognises is causation by psychological states.<sup>47</sup> His insistence that there be a connection between moral or normative requirements and states *inside* an agent is the aspect of his position that has been adopted by internalists. We see it, for example, in Williams' argument for internal reasons. Falk's position simply makes explicit the objection I raised against Williams' views: if we reject the possibility of normative motivation, then it is difficult to see how the account we give of an agent's behaviour can be an account of intentional action or action done for a reason. Falk's agent does his 'duty' because he is compelled to do so. The fact that he is compelled by causal forces inside his head rather than by external demands does not seem to alter the fact that he has no control over how he behaves.<sup>48</sup>

---

<sup>47</sup> Interestingly, Falk thinks that having a motive is not the same as having a desire. A motive is a thought that causes or causally implies an action. A desire is the 'feeling' of being compelled to do an action, or the very fact of being caused to do an action (Falk 1948: 117).

<sup>48</sup> Falk's view suggests that in order to be morally required to perform an act, an agent must not be free to do anything other than perform the act. This idea is opposed to ordinary notions of moral responsibility, according to which agents are held accountable for their actions precisely because they have freely chosen to perform them. (e.g. Korsgaard 1997: 244-247).

Despite these misgivings, I think it would be unwise to reject completely the intuition that generates the kind of internalism espoused by Falk and by Williams. There is a legitimate worry central to these views that moral requirements might be so removed from human experience and human capacities that they would have no relevance to and no hold on human behaviour (cf. Darwall 1995: 11; Korsgaard 1997: 239-245). In the next section I explore the possibility of extending my approach to moral judgement internalism to the issue of the connection between moral requirements and motivation.

### **8. Internalism about Moral Requirements and Motivation**

My principal concern in this work has been the connection between moral judgements and motivation. I have argued that moral judgement internalism is usefully thought of as an expression of our interested or participatory relation to moral considerations and moral practice. I now want to explore the possibility of extending this notion of participation to account for the purported internal connection between moral requirements and motivation. Obviously, I cannot deal fully with the issue of moral requirement internalism in the present work. What I aim to do is apply aspects of my approach to moral judgements to the issue of the relation between moral requirements and motivation. I suggest that if my account of moral judgements can be plausibly extended in a way that sheds light on internalist intuitions about moral requirements, then this will provide further support for my general approach.

I said in chapter one that the view that moral requirements entail motivation can be interpreted in a number of ways. Moral knowledge internalism says that motivation is in no way intrinsic to ethical facts but that being motivated is a necessary consequence of knowing ethical facts. By contrast, content internalism about moral requirements and motivation says that part of what makes it true that an agent is morally required to perform an act is that the agent has, or would have under suitable

conditions, some motivation to perform the act. According to content internalism, motivation is partly constitutive of moral requirements.<sup>49</sup>

The problem with moral knowledge internalism and the view of moral reasons it suggests is that it makes the ability of such reasons to motivate agents somewhat mysterious.<sup>50</sup> By contrast, content internalism is thought to secure a close and intelligible connection between a consideration's being a reason for and its capacity to motivate an agent (Darwall 1983: 52, 128).<sup>51</sup> Content internalism says that part of what an agent judges in making a moral judgement is that she is, or would be under suitable conditions, motivated to act in accordance with the judgement.<sup>52</sup> The presence of a motivating state in the content of a normative judgement is meant to explain how normative considerations can motivate. However, we saw in the discussion of Williams' view of internal reasons that content internalism fails to show how justifying considerations are capable of motivating intentional action.<sup>53</sup> Another problem with content internalism is that it contradicts the intuitively powerful Kantian doctrine that moral requirements apply to agents independently of their particular desires. These considerations suggest that a plausible account of the connection between moral requirements and motivation will not be a version of content internalism.

---

<sup>49</sup> A crucial difference between content and moral knowledge internalism, then, is that moral knowledge internalism takes moral reasons as metaphysically primary: the reason explains the motivation; whereas content internalism takes motivations as metaphysically primary: part of the explanation of an agent's having a reason or requirement to act in a certain way is that he has, or would have under certain conditions, some motivation to act in that way. This means that the direction of metaphysical dependence between normative facts and motivation differs between the two doctrines, even though both positions agree that having a reason to perform an act entails, under certain conditions, having a motivation to perform the act. Susan Hurley (2001) stresses the importance of distinguishing between the issue of direction of metaphysical dependence between reasons and motivation and the issue of logical relations between reason claims and motivation claims when discussing internalism. She argues that this distinction is obscured by Williams' account of internal reasons.

<sup>50</sup> Mackie expresses this worry (1977: 38).

<sup>51</sup> I stress again that I am assuming here that moral requirements provide normative reasons for action, and will use 'reasons', 'moral reasons', and 'moral requirements' interchangeably.

<sup>52</sup> Cf. Darwall (1983: 52. My emphasis): 'When we judge that some consideration is a reason... *part of what we judge* is a condition of our will.'

<sup>53</sup> The point being that positing a motivating state in the content of such judgements – content internalism – is meant to explain how justifying considerations are capable of motivating intentional action. In failing to do this, content internalism fails to achieve one of its principal aims.

If we are to develop an acceptable version of the thesis that there is an internal connection between moral requirements and motivation then our approach will have to avoid the extremes of moral knowledge internalism and of content internalism. To this end, I want to consider the very general intuitions that motivate internalism in ethics. Darwall claims that internalism, broadly construed, is the view that the desire to act for reasons must be ‘integral to the self’ rather than something ‘superadded to our nature’ (1983: 57).<sup>54</sup> On this view, internalism posits some form of intimate connection between an individual’s tendency to act for reasons and her agency. Moral requirement internalism suggests that the status of reasons or requirements themselves is connected to the kind of agency that disposes individuals to act for reasons. These considerations suggest the following interpretation of moral requirement internalism: moral requirements or reasons must be capable of motivating those to whom they apply, and this capacity must derive from an intimate connection between the nature of such requirements or reasons and the natures of the individuals to whom they apply.

I suggest that this interpretation of moral requirement internalism can be accommodated by the model of moral agency and practice that grounds my account of moral judgements.<sup>55</sup> Consider the internalist claim that the desire to act for reasons must be in some sense integral to the nature or identity of an agent. This claim can be captured by my notion of the normatively engaged agent. Recall that an individual is normatively engaged just in case she is generally susceptible to the range of reactive attitudes that make up the participant point of view. Being normatively engaged in this manner is a condition of normative competence: it enables agents to hold themselves and others to expectations and thus enables them to be guided by moral considerations. This means that normative competence is

---

<sup>54</sup> Cf. Korsgaard: [T]he dogmatic rationalist [one who sees reasons as external to the agent] is unable to explain how reasons get a grip on the agent, because he supposes that reasons exist independently of the rational will... This model... seems to invite the question: but suppose I don’t care about being rational? What then?’ (1997: 243-244).

<sup>55</sup> The general form of the following argument is similar to Velleman’s approach to the internal/external reasons debate in *The Possibility of Practical Reason* (1996). Velleman argues that the debate about whether normative reasons are or are not in part constituted by agents’ motivations embodies a false dichotomy and suggests an alternative position that aims to capture both internalist and externalist intuitions.

partly constitutive of the status of individuals as moral agents. The tendency to be guided by normative considerations is therefore integral to the nature of such agents, as required by internalism.

Next, consider the internalist claim that the capacity of individuals to be motivated by normative requirements must derive from a connection between the nature of those *requirements* and the natures of the individuals to whom they apply. This condition is met by my account of the connection between normative engagement and normative requirements. Recall that severe normative incompetence not only means an individual cannot make genuine moral judgements, it means *she is no longer an appropriate object of moral or normative appraisal*. So the *status* of a consideration as a requirement or reason for an individual is dependent upon the individual's being appropriately normatively engaged. And being normatively engaged in this way is just what explains the capacity of agents to be motivated by normative considerations. So the model of normative engagement and normative competence satisfies the various aspects of our revised version of moral requirement internalism.

This, then, is a very general outline of how the approach to moral judgements I have advocated might be extended to accommodate internalist intuitions about moral requirements. If such a move is plausible, then I think it provides further support for the general approach to moral judgements and moral motivation that I have been considering in this thesis.

## **9. Conclusion**

The speech act approach to moral judgement must establish a suitable connection between moral thought and moral language to explain how linguistic rules can have implications for moral psychology. I argue that the human capacity for language plays a central role in the inherently social practice of moral evaluation. The norms constitutive of this practice are partly structured by a presupposition of interest in moral deliberation. To a large degree evaluative practice is realised linguistically, and the interest in moral deliberation is encoded in the rules governing the



performance of moral speech acts. The picture of moral motivation suggested by the speech act account cannot easily be accommodated by the desire/belief model of motivation and action explanation. While the standard Humean distinction between cognitive and conative states seems correct, it is not clear that the sense of 'motivation' essential to conative states is the only sense of 'motivation' used in the explanation of intentional action. We should distinguish between normative motivation and conative motivation, particularly given the crucial role the concept of normative motivation plays in our ordinary notion of intentional action.

Internalists argue that there needs to be a substantial connection between moral considerations and the individuals to whom such considerations apply in order to secure the justifying authority and motivating force of morality. The notions of normative competence and normative engagement, central to my account of moral judgement, can be extended to provide a plausible account of this connection.

## Conclusion

---

In an early and influential discussion of the debate between internalists and externalists, Frankena suggests the following compromise position:

He [the externalist] may hold...that judgements of obligation have a conceptual content of an "external" kind, but add that we do not speak of a man's *assenting* or *sincerely assenting* to them unless he not only apprehends the truth of their conceptual content but is at least to some extent moved to conform to it. He would then admit that it is *logically* possible that one might have a "mere intellectual apprehension"...of their truth, but he would recognise the generally practical function of language..., especially moral discourse. (Frankena 1958: 66)

To some extent my approach in this thesis can be seen as a development of Frankena's suggestion. I have argued that we should reject the assumption that internalism provides a *prima facie* obstacle to cognitivism in ethics. I have suggested that by distinguishing between the act and content of moral judgements we can secure a necessary connection between moral judgements and motivation while respecting the independence of moral requirements from agents' contingent desires and inclinations. The speech act approach provides a structured way of combining the practical and motivating aspects of moral judgements with the cognitive and fact-stating aspects of such judgements. It accommodates the view that moral judgements entail endorsing attitudes, a version of moral judgement internalism, and explains cases in which, due to depression or accidie, moral judgements fail to motivate. It does this by stressing the practical function of moral language as embodied in the conditions for the successful and sincere performance of moral speech acts.

The speech act analysis of moral judgements also sheds light on the nature of moral psychology and of moral motivation. It presents, I think, a plausible explanation of the sense in which, in Hume's terms, morality goes 'beyond the calm and indolent judgements of the understanding' (1739-40: Book III. Part I. Sect. I). In making moral judgements, agents commit themselves to acting in accordance with judged requirements. This commitment, we might say, is an agent's practical recognition of the normative authority of the requirement she asserts. The importance of this

commitment is suggested by the readiness with which we question someone's moral sincerity if the commitment is absent. This intuition is perplexing if, as cognitivists, we hold that sincerity only requires belief. It reflects, I think, the central role that the notion of holding responsible or accountable plays in our ordinary understanding of what it is to make a moral judgement.

It is important to understand the role of moral judgement in broader moral practice. Moral judgement is properly defined by the role it has in this practice, including its connection to moral attitudes and responses, and its relation to notions of moral agency and accountability. However, there are implications of this position with which I am not entirely comfortable. On one reading of normative competence, suggested by certain aspects of Strawson's position, agents who fail to recognise the demands that others make on them are classed as normatively incompetent and, as such, outside the realm of moral accountability and moral engagement. I take it that a plausible conception of normative competence must be able to accommodate moral views that might be offensive or significantly removed from one's moral perspective. Otherwise we may find ourselves adopting the objective stance towards anyone with whom we significantly disagree.

These issues are linked to the nature of the reactive and sanctioning attitudes that make up the participant point of view. If the development of such higher-order attitudes is fundamentally linked to human sociality and language, it seems we run the risk that different cultural and social circumstances will produce different, perhaps incommensurable, sets of reactive attitudes. The worry is that by removing the internalist obstacle to cognitivism, and by extension to realism, we have created a different kind of obstacle to objectivism and realism in ethics. Nonetheless, I do think that any point of view that is recognisably moral, in the broad sense in which I have used this term, will have to incorporate some conception of normative motivation and with it some conception of responsible agency. It is perhaps in these and related notions that we will find the basis of a point of view with which all human agents, by their very nature, are engaged.

## Bibliography

---

- Altham, J.E.J., and Harrison, R. (eds.) 1995. *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press.
- Austin, J.L. 1962. *How To Do Things With Words*. Oxford: Oxford University Press.
- Ayer, A.J. 1946. *Language, Truth, and Logic*. 2<sup>nd</sup> ed. London: Gollancz.
- Baron, M. 1984. "The Alleged Moral Repugnance of Acting from Duty." *The Journal of Philosophy* 81: 197-220.
- Berlin, I. 1973. *Essays on J.L. Austin*. Oxford: Oxford University Press.
- Blackburn, S. 1984. *Spreading the Word*. Oxford: Oxford University Press.
- Boyd, R. 1988. "How to be a Moral Realist." In Sayre-McCord 1988: 181-228.
- Brink, D. 1989 *Moral Realism and the Foundation of Ethics*. Cambridge: Cambridge University Press.
- 1997 "Moral Motivation." *Ethics* 108: 4-32.
- Copp, D. 1995. "Moral Obligation and Moral Motivation." *Canadian Journal of Philosophy* suppl. vol. 21: 187-219.
1997. "Belief, Reason, and Motivation." *Ethics* 108: 33-54.
2001. "Realist Expressivism: A Neglected Option for Moral Realism." In Paul, Miller, and Paul 2001: 1-43.
- Cullity, G. and Gaut, B. (eds.) 1997. *Ethics and Practical Reason*. Oxford: Clarendon Press.
- Dancy, J. 1993. *Moral Reasons*. Oxford: Blackwell.
2000. (ed.) *Normativity*. Oxford: Blackwell.
2000. *Practical Reality*. Oxford: Oxford University Press.
- Darwall, S. 1983. *Impartial Reason*. Ithaca, N.Y.: Cornell University Press.
1995. *The British Moralists and the Internal 'Ought': 1640-1740*. Cambridge: Cambridge University Press.
- Davidson, D. 1980 *Essays on Actions and Events*. Oxford: Oxford University Press.
- Dreier, J. 2000. "Dispositions and Fetishes: Externalist Models of Moral Motivation." *Philosophy and Phenomenological Research* 61: 619-638.
- Evans, D. 2001. *Emotion*. Oxford: Oxford University Press.
- Falk, W.D. 1948 " 'Ought' and Motivation." *Proceedings of the Aristotelian Society* 48: 111-138.
- Fischer, J.M., and Ravizza, M. 1993. *Perspectives on Moral Responsibility*. Ithaca, N.Y.: Cornell University Press.
- Foot, P. 1972. "Morality as a System of Hypothetical Imperatives." Reprinted in Foot 1978: 157-173.
1978. *Virtues and Vices*. Los Angeles: University of California Press.

- Forguesson, L.W. 1973. "Locutionary and Illocutionary Acts." In Berlin 1973: 160-185.
- Fotion, N. 2000. *John Searle*. Teddington: Acumen.
- Frankena, W.K. 1958 "Obligation and Motivation in Recent Moral Philosophy." In Meldon 1958: 41-81.
- Geach, P. 1958. "Assertion." *Philosophical Review* 74: 449-465.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgement*. Cambridge, Mass.: Harvard University Press.
- Grice, H.P. 1971. "Intention and Uncertainty." *Proceedings of the British Academy* 57: 263-279.
- Griffiths, P. 1997. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press.
- Hare, R.M. 1952. *The Language of Morals*. Oxford: Oxford University Press.  
1981. *Moral Thinking: Its Levels, Method and Point*. Oxford: Oxford University Press.
- Harman, G. 1975. "Moral Relativism Defended." *Philosophical Review* 85:3-22.
- Hooker, B. 1987. "Williams' Argument Against External Reasons." *Analysis* 47: 42-44.
- Humberstone, L. 1992. "Direction of Fit." *Mind* 101: 59-83.
- Hume, D. 1739-40. *A Treatise of Human Nature*. (ed.) E.C. Mossner. London: Penguin, 1969.
- Hurley, S.L. 1989. *Natural Reasons*. New York: Oxford University Press.  
2001. "Reason and Motivation: The Wrong Distinction?" *Analysis* 61: 151-155.
- Johnston, M. 1989. "Dispositional Theories of Value – pt. III." *Proceedings of the Aristotelian Society*, supp. vol. 63: 139-174.
- Korsgaard, C.M. 1986. "Scepticism About Practical Reason." *The Journal of Philosophy* 83: 5-25.  
1997. "The Normativity of Instrumental Reason." In Cullity and Gaut 1997: 215-254.
- Langton, R. 1993. "Speech Acts and Unspeakable Acts." *Philosophy and Public Affairs* 22: 293-330.
- Lewis, D. 1988. "Desire as Belief." *Mind* 97: 323-332.  
1989. "Dispositional Theories of Value – pt. II." *Proceedings of the Aristotelian Society*, supp. vol. 63: 113-137.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
- McDowell, J. 1978. "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society* supp. vol.52: 13-29.  
1979. "Virtue and Reason." *The Monist* 62.: 331-50.  
1985. "Values and Secondary Qualities." Reprinted in Sayre-McCord 1988: 166-180.

1995. "Might there be External Reasons?" In Altham and Harrison 1995: 68-85.
- Meldon, A.I. 1958. (ed.) *Essays in Moral Philosophy*. Seattle: University of Washington Press.
- Mele, A.R. 1996. "Internalist Moral Cognitivism and Listlessness." *Ethics* 106: 727-753.
- Nagel, T. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.
- Paprzycka, K. 1999. "Normative Expectations, Intentions, and Beliefs." *The Southern Journal of Philosophy* 37: 629-652.
- Parfit, D. 1997. "Reasons and Motivation." *Proceedings of the Aristotelian Society* suppl. vol. 71: 99-130.
- Paul, E.F., Miller, F.D., Paul, J. 2001. *Moral Knowledge*. Cambridge: Cambridge University Press.
- Price, H. 1988. "Defending Desire-as-Belief." *Mind* 97: 119-127.
- Railton, P. 1986. "Moral Realism." *Philosophical Review* 95 : 163-207.
2000. "Normative Force and Normative Freedom." In Dancy 2000: 1-33.
- Rawls, J. 1973. *A Theory of Justice*. Oxford: Oxford University Press.
- Ridge, M. 2002. "Acceptance, Sincerity and Expressivism" (unpublished draft)
2004. "Sincerity and Expressivism" (unpublished draft)
- Sayre-McCord, G. (ed.) 1988. *Essays on Moral Realism*. Ithaca: Cornell University Press.
- Schueler, G.F. 1995. *Desire: Its Role in Practical Reasoning and the Explanation of Action*. Cambridge, Mass.: MIT Press.
- Searle, J. 1968. "Austin on Locutionary and Illocutionary Acts." In Berlin 1973: 141-159.
1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
1979. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
1989. "How Performatives Work." *Linguistics and Philosophy* 12: 535-558.
1995. *The Construction of Social Reality*. London: Allen Lane, The Penguin Press.
1998. *Mind, Language and Society: Philosophy in the Real World*. New York: Basic Books.
- Shafer-Landau, R. 2000. "A Defense of Motivational Externalism." *Philosophical Studies* 97: 267-291.
- Simpson, E. 1999. "Between Internalism and Externalism in Ethics." *The Philosophical Quarterly* 49: 201-214.
- Sinnott-Armstrong, W. 1999. "An Argument for Descriptivism." *The Southern Journal of Philosophy* 37: 281-91.
- Smith, M. and Pettit, P. 1990. "Backgrounding Desire." *The Philosophical Review* 99: 565-592.
- Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell.



1997. "In Defense of 'The Moral Problem'." *Ethics* 108: 84-119.
- Stevenson, C.L. 1937. "The Emotive Meaning of Ethical Terms." *Mind* 46: 14-31.
1944. *Ethics and Language*. New Haven, Conn.: Yale University Press.
- Stocker, M. 1979. "Desiring the Bad: An Essay in Moral Psychology." *Journal of Philosophy*: 738-53.
- Strawson, G. 1986. "On 'Freedom and Resentment'." Reprinted in Fischer and Ravizza 1993: 67-100.
- Strawson, P. 1962. "Freedom and Resentment." Reprinted in Fischer and Ravizza 1993: 45-66.
- Svavarsdóttir, S. 1999. "Moral Cognitivism and Motivation." *The Philosophical Review* 108: 161-219.
- Velleman, J. David. 1989. *Practical Reflection*. Princeton, N.J.: Princeton University Press.
1992. "The Guise of the Good." *Nous* 26, 1: 3-26.
1996. "The Possibility of Practical Reason." *Ethics* 106: 694-726.
- Wallace, R.J. 1994. *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press.
- Watson, 1987. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." Reprinted in Fischer and Ravizza 1993: 119-148.
- Williams, B. 1976. "Persons, Character and Morality." Reprinted in Williams 1981: 101-13.
1980. "Internal and External Reasons." Reprinted in Williams 1981: 101-113.
1981. *Moral Luck*. Cambridge: Cambridge University Press.
1985. *Ethics and the Limits of Philosophy*. Cambridge, Massachusetts: Harvard University Press.
1989. "Internal Reasons and the Obscurity of Blame." Reprinted in Williams 1995: 35-45.
1995. *Making Sense of Humanity*. Cambridge: Cambridge University Press.
- Wright, C. 1988. "Moral Values, Projection and Secondary Qualities." *Proceedings of the Aristotelian Society* supp. vol 62: 1-26.
1995. "Truth in Ethics." *Ratio* 8: 209-226.