THREE ACCOUNTS OF COGNITIVIST INTERNALISM UNDERMINED

by

Anthony Skelton

Submitted in partial fulfillment of the requirement
for the degree of Master of Arts

at

Dalhousie University
Halifax, Nova Scotia
August, 1997

Table of Contents

Chapter One **Two Intuitions and Hume's Dictum: The Moral Problem**

Introduction 1

1.1 Two Intuitions 2

1.2 Hume's Dictum 6

1.3 The Moral Problem 9

1.4 Putative Solutions to the Moral Problem 14

1.5 Conclusion 18

Chapter Two **David McNaughton: Internalist Moral Realism**

Introduction 19

2.1 Categorical and Hypothetical Imperatives 19

2.2 A Cognitive Account of Motivation 29

2.3 A Cognitive Account of Motivation Undermined 33

2.4 The Case of the Masochist 37

2.5 Conclusion 42

Chapter Three **Jonathan Dancy: The Pure Theory**

Introduction 44

3.1 The Pure Theory 44

3.2 Amoralism, Wickedness and Lucifer Himself 47

3.3 Weakness of Will and the Person Suffering from Accidie 58

3.4 Particularism Undermined 65

3.5 The Argument from Directions of Fit 71

3.6 Further Problems for the Pure Theory 79

3.7 Conclusion 83

Chapter Four **Michael Smith: Cognitivism, Internalism and Humeanism At Once**

Introduction 85

4.1 The Practicality Requirement 86

4.2 Rationalism 98

4.3 Smith's Account of Normative Reasons 103

4.4 Smith's Account of Normative Reasons Undermined 108

4.5 Conclusion 114

Bibliography 115

Abstract

Consider the following intuitively plausible, though mutually inconsistent propositions:

1. Moral judgements, e.g., "It is right to Φ", express only an agent's beliefs about an objective matter of fact, and not her desires.

2. If an agent judges that it is right to Φ, then, *ceteris paribus*, she is motivated to Φ.

3. An agent is motivated to act just if she has a desire and a means-end belief, where belief and desire are distinct existences.

The inconsistency arises because 1 states that a moral judgement expresses a belief, which, according to 2, motivates an agent to act, which must be a desire according to 3. Thus, 1, 2 and 3 together entail that there is a necessary connection between having a belief and having a desire. But 3 denies that there is such a connection, given that beliefs and desires are distinct existences. This is the moral problem. In the first chapter, I explain the nature of this problem. In the remainder of the essay, I canvass and reject three recent solutions to the moral problem. In chapters two and three I discuss David McNaughton's and Jonathan Dancy's attempts to solve this problem by rejecting 3. I argue that they fail to adequately defend their own positive views and fail to rebut 3. In the final chapter of the essay, I discuss Michael Smith's solution to the moral problem. Smith attempts to solve the problem by reconciling the statements. I argue that he fails to defend 2 and that he cannot make 2 and 3 consistent.

Acknowledgements

An immense amount of gratitude goes to Duncan MacIntosh, my supervisor, for his helpful comments and criticisms. Thanks to Richmond Campbell and Nathan Brett for reading this essay. I would also like to thank the entire faculty for making this year one of the most intellectually challenging and rewarding years of my life. I own a great intellectual debt to Margaret Cameron who has read, critiqued and commented on everything I have written this past year. Thanks to my parents, Marie and Peter Skelton, for their moral and financial support.

CHAPTER ONE

Two Intuitions and Hume's Dictum: The Moral Problem

'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger.
- David Hume[1]

Moral philosophy is a cake with three layers. The first layer is practical or 'applied' ethics, which deals with substantive moral problems such as whether it is wrong to have an abortion, or whether the unbridled pursuit of economic gain is a permissible business practice, or whether the government's taxation proposal is unjust. The second layer comprises ethical theory, which concerns itself with construction of a general theory of what makes an action right or wrong, or a person good or bad. These theories are employed (or 'applied') when answering specific practical moral questions.

Finally, there is metaethics. Metaethics deals with a broad range of issues concerning the nature of moral judgements. According to Stephen Darwell, Allan Gibbard and Peter Railton[2], "metaethics studies . . . the justification and justifiability of ethical claims as well as their meaning, and also the metaphysics and epistemology of morals, and like matters." Rather than dealing with answers to questions like "Should I lie?", metaethics deals with questions *about* these types of questions. For example, does a moral judgement such as "Stealing is wrong," state a fact? Can moral judgements be true? Can we justify our moral judgements? Is there one true moral theory? What does the 'should' in "Should I lie?" mean? Does the fact that I recognize that I have a moral

[1] Hume 1978, 416.
[2] Darwell, Gibbard and Railton 1992, 125-26.

obligation necessarily motivate me to act in accordance with what the obligation enjoins?[3] The questions with which I will deal in this essay are metaethical questions.

1.1 Two Intuitions

According to Michael Smith[4], metaethical debate engenders radical disagreement. Smith suggests that this disagreement can be traced to two distinct features of morality, namely, its putative objectivity and its practicality (or its action-guiding nature). It is important to notice here that these features are not merely figments of a philosophically over-inventive mind. Nor are they merely theoretical. Rather, they are embedded in common-sense moral thinking, and, hence, should best be thought of as found in intuitions about morality itself. What is paradoxical about these intuitions is that they "pull against each other, so threatening to make the very idea of morality incoherent."[5] This problem, which besets contemporary metaethics, Smith dubs the 'Moral Problem'. In the remainder of this chapter, I will describe the moral problem as it is presented by Smith[6], and then I will recount several attempts by others to avoid it.

The moral problem, then, starts with two intuitions or features of morality. According to Smith, the first is this: when we engage in moralizing with the goal of answering certain moral questions, e.g., 'Should I lie to spare my friend's feelings?', we are concerned to come, not to the answer which best suits our own needs, but to the 'correct' or 'right' answer. And, this presupposes that there *can* be 'correct' or 'right'

---

[3] Although I have divided moral philosophy up into three layers, there are several alternative ways of carving up this territory. For example, David Brink argues that the metaethical issues are second-order ethical issues, and that both practical ethics and ethical theory are first-order levels of inquiry in ethics. See Brink 1989, 1-2. For similar terminology see Mackie 1977, 16.

[4] Smith 1994, 4..

[5] Smith 1994, 5. See also Smith 1991, 399.

[6] Smith 1994, 1-15.

answers to moral questions. In a similar vein, David Brink argues that when we engage in a moral argument or discussion, "we try and hope to *arrive* at the correct answer, that is, the answer that is correct prior to, and independently of, our coming upon it."[7] This presupposes that there exists a mind-independent moral reality, one which contains moral facts which determine the rightness of moral judgements and about which we can form beliefs; and which, in virtue of their mind-independent status, we can be wrong about.

Smith contends that the intuition that there are such facts explains why we care about and engage in moral arguments at all, and why we agonize over moral questions. We do these things because we wonder if the answer we give to a moral question is the right one. If we didn't think there was a right answer, then why would we argue about it or wonder if the view we hold is the right one? Brink concurs that our everyday moralizing presupposes that there are correct answers to moral questions. According to Brink, "when we are uncertain about moral issues, we often deliberate as if there were a right answer to the issue before us."[8] The fact that we intuitively feel that there right answers to moral questions seems to explain why we engage in moralizing at all.

Furthermore, Smith argues, we intuitively feel that the only thing which determines the 'rightness' of an action (or the moral fact of the matter about the situation) is the circumstance in which it takes place, not things like one's own feelings. In other words, the non-moral facts which comprise the circumstance in which the action takes place 'fix' or 'fashion' the moral properties of the action, so that if any two people carry out the same action in the very same circumstances (i.e., if all the non-moral facts about

---

[7] Brink 1989, 31.
[8] Brink 1989, 29.

them and their actions are the same), then their actions have the very same moral property, e.g., rightness or wrongness.

Smith says that we seem to think (or intuit) that the best way to access or discover what the moral facts in a situation are is to formulate theories and marshal reasons in support of or against any answer to a moral question, dilemma or issue. Moreover, if we are willing to listen to each other and take each other's positions and view-points seriously, then, we seem to hold, the result of the marshaling of reasons pro and con will be "a convergence on moral opinion, a convergence on the truth."[9]

> This first feature or intuition about ethics can be summed up as follows:
>
> We seem to think that moral questions have correct answers; that the correct answers are made correct by objective moral facts; that moral facts are wholly determined by circumstances; and that, by engaging in moral conversation and argument, we can discover what these objective facts determined by the circumstances are.[10]

Smith calls the first feature of morality the "objectivity of moral judgement."[11] Moral judgements, then, purport to be objective: they are made correct or incorrect by the existence of moral facts which are determined by the circumstances in which the judgements occur; and they are discovered through earnest moralizing.

According to Smith, the second feature of morality concerns the "practicality of moral judgement."[12] The best way to illustrate this feature of morality is by way of an example of Smith's: suppose that after a long conversation debating the pros and cons of

---

[9] Smith 1994, 6.
[10] Smith 1994, 6.
[11] Smith 1994, 6.
[12] Smith 1994, 6.

donating to famine relief I have been convinced, by my interlocutor, that I have this obligation. That is, I believe that I have an obligation to donate to famine relief. It seems plausible to assume of me that, given that I believe that I have an obligation that I should give to famine relief, I will give to famine relief when I have the opportunity to do so. However, suppose that, despite the fact that I believe that I should give to famine relief, I claim, when presented with the next opportunity to act in accord with my obligation, that though I have been convinced that I *should* give to famine relief, I have not been convinced that I have a *reason* to give to famine relief.

My reaction, Smith tells us, would give rise to puzzlement. This follows from the fact that if I am convinced (i.e., I believe) that I should give to famine relief, then it also seems to be the case that I am convinced that I *do have reason* to give to famine relief. That is, the conversation I had with my interlocutor about whether or not I should give to famine relief seems "equivalent to a conversation about whether or not I have *reason* to give to famine relief."[13] To say that I should give to famine relief just is to say that I have reason to give to famine relief. Furthermore, reasons motivate people to act: we take it for granted that when an agent acts intentionally, she has reasons for doing what she does. Now, if I accept that I have reason to give to famine relief, then I have a motive for giving to famine relief. And, in the absence of contrary motives, if I have a motive, I will be induced to act. Smith concludes from this that, if I am convinced that I have reason to give to famine relief, and reasons furnish me with a motive, and motives produce action,

[13] Smith 1994, 6; emphasis added.

then, absent weakness of will or other psychological failings, it would seem quite odd that I would not be moved to act by the reasons I have.

The second feature of morality can be summarized as follows:

> Moral judgements seem to be, or imply, opinions about the reasons we have for behaving in certain ways, and, other things being equal, having such opinions is a matter of finding ourselves with a corresponding motivation to act.[14]

1.2 Hume's Dictum

To make sense of this second feature of morality, we must (following Smith) digress somewhat to discuss more broadly the nature of human psychology and action explanation. The received, though not undisputed,[15] theory of human psychology is given to us by David Hume.[16] According to Hume, there are two distinct types of mental states: beliefs and desires. Beliefs attempt to represent the world as it is. As such, they can be assessed as either true or false according to whether they accurately or inaccurately represent the world. Desires, on the other hand, are not representational, they do not attempt to represent the world as it is. Consequently, they cannot be assessed for truth or falsity. Desires are not states which merely represent the world, they the active states of an individual which move her to change the world. Thus, according to Hume, beliefs and desires are wholly different.

There is a further way in which beliefs and desires are different. Whereas beliefs, since they are representational, can be rationally criticized, desires, since they are not

---

[14] Smith 1994, 7.
[15] Nagel 1970 and Platts 1979.
[16] Hume 1978, 413-418, 455-459.

representational, cannot be rationally criticized; that we have desires is merely a fact about ourselves. Consequently, changing or altering our beliefs about the world in the face of newly acquired evidence does nothing to affect, at least rationally, the desires we hold, even though it may change our beliefs about the world. [17]

Another way of making the difference between beliefs and desires clear is presented to us by Mark Platts.[18] According to Platts, what marks the distinction between beliefs and desires is that they have different and opposing 'directions of fit'. Beliefs track the truth: if they are true, then they fit the world; if they are false, they do not fit the world, and hence, should be jettisoned. "False beliefs should be changed to fit the world and not vice versa."[19] On the other hand, desires aim at their own satisfaction: when the target conditions of a desire obtain, the world fits it. However, if the target conditions of a desire do not obtain, it need not, unlike a belief, be jettisoned. As Platts explains: "the world . . . should be changed to fit with our desires, not vice versa."[20] Thus, the idea that beliefs and desires have differing and opposing directions of fit points to the fact that they are radically different kinds of states.

Before going forward with our discussion of the second feature of morality, it is important to head off one misconception regarding Humean desires. Suppose that I hold a desire which is based on a false belief, e.g., I desire never to be out past midnight because

---

[17] It is important to point out here a salient difference between the contemporary view regarding desires and Hume's account of desires. According to the contemporary account, desires are representational, just not in the same way that beliefs are. Desires represent the way the world is to be once their target conditions obtain. This, however, does not change the fact that desires are immune to rational criticism, since they do not purport to represent the world as it is.

[18] Platts 1979, 256-7.

[19] Platts 1979, 256.

[20] Platts 1979, 257.

I believe that if I stay out past midnight, I will turn into a pumpkin. In this case, I have an 'irrational' desire. Now, my present desire, the desire not to go out past midnight, is founded on a further desire and a belief: my desire not to turn into a pumpkin and my belief that if I stay out past midnight, then I will turn into a pumpkin. It follows from Hume's theory above that I can be rationally criticized for holding the belief I do by virtue of the fact that it is false. Since my belief is false, and, since my desires are based on that false belief, I can be rationally criticized for holding the desire not to be out past midnight. Thus, desires can be rationally criticized if they exist by virtue of the holding of a false, hence rationally criticizable, belief. However, desires not so generated, i.e., not produced by false beliefs, are not rationally criticizable.

Smith contends that Hume's theory of human psychology is relevant to the second feature of morality because it provides us with a framework within which we can explain intentional human action. Beliefs represent how the world is, and hence, how the world can be appropriately altered to the end of making it the way the target conditions of our desires tell us it should be. As Smith puts it, "an action is thus the product of these two distinct existences: a desire representing the way the world is to be and a belief telling us how the world is to be changed so as to make it that way."[21] That desires and beliefs have distinct existences and that an agent is motivated just if she possesses a desire and an appropriate means-ends belief is what I will call Hume's dictum.

---

[21] Smith 1994, 9.

1.3 The Moral Problem

Returning now to the first feature of morality, the objectivity of morality, we can see that by virtue of holding the intuitions in the first feature one is naturally led to certain philosophical positions. The first of these is moral realism. According to David Brink, moral realism, roughly speaking, is the metaphysical view that "there are moral facts and true moral claims whose existence and nature are independent of our beliefs about what is right and wrong."[22] Those who affirm the intuitions comprising the first feature of morality hold that there are correct answers to moral questions and that these answers are made correct by virtue of objective moral facts, i.e., facts that determine the rightness and wrongness of actions, or the goodness or badness of persons. Hence, they are naturally led, it seems, to embrace moral realism. Now, it is also true about the first feature of morality that it supposes that we can have beliefs about what is right and wrong, i.e., about moral facts, that moral knowledge is possible and that moral judgements express beliefs. To think that we can form beliefs about these moral facts, and, hence, that moral judgements express beliefs about the way things are morally, is to embrace cognitivism. This view is called cognitivism because it says that we can form beliefs, representational cognitive states, about what morally is the case, i.e., we can have representations of the way the world is, morally speaking.

There is a further view about moral motivation which, given the Humean theory of action explanation outlined above, is the natural concomitant of those who espouse moral realism, cognitivism and Hume's dictum, namely, the view that the mere holding

---

[22] Brink 1989, 7.

of a moral belief will not by itself motivate a person to act in accordance with what she believes is right. This view is called externalism because it maintains that believing that one has a moral obligation is insufficient to motivate one to act; something external to the moral judgement must be the source of motivation. Brink explains that externalism is the view "that whether moral considerations motivate or provide reason for action depends on factors external to the concept of morality, such as the content of morality, a substantive theory of reasons for action, and facts about the world, such as the agent's interests or desires."[23] One may desire, or be disposed to desire, to do what one believes morality enjoins; or one may not; it is merely a contingent fact about moral obligations that they motivate. If one does not desire, or is not disposed to desire, to do what morality enjoins, and one's desire does not rest on a false moral belief, then one cannot be rationally criticized, since desires which do not rest on false beliefs are immune from rational criticism. Thus, the fact that I morally should do Φ *does not* provide me with *reason* to do Φ.

Holding the intuition espoused by the first feature of morality naturally leads to jettisoning the intuitions which arise in a discussion of the second feature of morality. Those who are moral realists, and who at the same time accept Hume's dictum, must espouse externalism, and so they must deny that a belief that one may have that one has a moral obligation will motivate one to act in accord with that obligation. Thus, given realism and Hume's dictum, it is not necessarily the case that if I judge it right to give to famine relief, I will be motivated to give to famine relief.

---

[23] Brink 1989, 8.

The second feature of morality, that to have a moral opinion is, *ceteris paribus,* to find oneself with a corresponding reason or motivation to act, leads to a series of philosophical positions. If Hume is right that only desires conjoined with appropriate beliefs (and not beliefs themselves) motivate one to act, and if, as the second feature of morality states, when one has a moral opinion, one has a reason (or is motivated) to act, then moral opinions or judgements must express desires. Smith explains that the psychological implication of this feature of morality is this: "since making a moral judgement requires our having a certain desire, and no recognition of a fact about the world could rationally compel us to have one desire rather than another, our judgement must really simply be an expression of that desire, or perhaps a complicated disposition to have that desire."[24] The view that moral judgements are simply and solely expressions of a desire is called non-cognitivism.

Smith argues that the second feature of morality also has a metaphysical implication, namely, irrealism. Irrealism, or anti-realism, is the view that when we judge that something, Φ, is wrong, we are merely expressing our desire that Φ cease; we are not, contra cognitivism, expressing a belief. That is, when we make a moral judgement we are not stating a fact about the way things are morally; in fact, we are not mentioning facts at all. Thus, irrealism about ethics denies that there are such things as moral facts.

Smith suggests that if we embrace this form of irrealism we can explain quite easily how it is that our moral judgements connect up neatly with our motivation to act in accordance with our moral obligations. Those who espouse irrealism or non-cognitivism

---

[24] Smith 1994, 10.

typically embrace internalism. According to Brink, internalism is the view that "there is an internal or conceptual connection between moral considerations and action or the sources of action."[25] When I make a moral judgement I am expressing a desire, and since desires, when conjoined with appropriate means-ends beliefs, will motivate me to act, I will be inclined to act in accordance with my moral judgement, i.e., my desire. Thus, the mere recognition that I have a moral obligation will motivate me to act in accord with that obligation. Furthermore, Smith argues, when I judge that it is right to give to famine relief, in addition to expressing a desire, "I commit myself to saying that it is right for anyone in circumstances like mine to give to famine relief as well."[26] But, Smith argues, it is not the case that I think that other people or moral agents therefore have reason to give to famine relief. *A fortiori,* I do not think that my claim that others who are in the same circumstances as mine should give to famine relief is in any way an objective truth. This follows from the fact that since a moral judgement in an expression of a desire, and a desire is not rationally criticizable unless it rests on a false belief, a moral judgement, if it does not rest on a false belief, is immune to rational criticism. The fact that others might disagree with my moral judgement is not a sign of their irrationality; it simply points to the fact that they have desires different from mine.

This, Smith explains, brings us to the moral problem. The first feature of morality can quite easily explain moral argumentation because it says that there is a fact of the matter about what is wrong, and that this fact can be discovered through genuine moralizing. These facts are determined by objective moral facts which, in turn, are

---

[25] Brink 1989, 38.
[26] Smith 1994, 10.

determined by the circumstances in which the judgement takes place. Furthermore, the first feature of morality alleges that moral judgements express beliefs. However, the first feature or morality cannot, given Hume's theory of action explanation, explain how the having of a moral opinion can motivate one to act. It cannot explain why it is that we expect people who make certain moral judgements to act in accord with their moral judgements. On the other hand, the second feature of morality says that moral judgements express desires, and so can easily explain the connection between the having of a moral opinion and being motivated to act in accordance with it. Yet, it cannot explain what we are arguing about when we engage in moral argument.

The moral problem arises, then, because a moral judgement cannot be everything that we intuitively take it to be. It cannot be simultaneously objectively correct or incorrect in the way of a belief *and* practical or action-guiding in the way of a desire. Nothing can both be a fact or a belief about a fact and, at the same time, given Hume's dictum, be action-guiding.

According to Smith, the moral problem explains the source of disagreement in the domain of metaethics; in fact, he alleges, "the moral problem is the central and organizing problem in contemporary metaethics."[27] These three inconsistent statements capture the essence of the moral problem:

> 1. Moral judgements of the form 'It is right that I Φ' express [only] a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.[28]

---

[27] Smith 1994, 11.

[28] I add the world "only" to Smith's reconstruction of 1, because without this addition it is not entirely clear that the three statements are inconsistent. On Smith's reconstruction, 1 does not rule out the possibility that a moral judgement could contain both an agent's beliefs and desires. If this were the case, then we could hold all three of the statements consistently, since 1 and 2 no longer seem in conflict with 3: an agent is motivated by a moral judgement, which expresses a desire in addition to a belief.

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

3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences.[29]

These statements are mutually inconsistent in the following way. 1 states that moral judgements express beliefs about the way things are morally speaking, while 2 states that there is an internal or conceptual connection between an agent's making a moral judgement (i.e., having a moral belief) and acting in accordance with what it enjoins. But, given Hume's dictum, 3, there is no such connection between believing that one has a moral obligation and acting in accordance with it, because facts, or beliefs about them, do not motivate one to act: Beliefs are motivationally inert, and only desires in conjunction with the appropriate means-end belief can motivate. Thus, the moral problem is that moral judgements cannot simultaneously be objective and practical.

1.4 Putative Solutions to the Moral Problem

Thus far I have sketched some of the views which arise out of the two features of morality and that comprise the main cluster of views in contemporary metaethics. As will become clear in this section, some of these very positions, namely, non-cognitivism, externalism and anti-Humeanism, purport to provide solutions to the moral problem. My goal in this section is to give a brief sketch of the various solutions to the moral problem.

I have not argued against the possibility that a moral judgement might express both a belief and a desire. For the purposes of this essay I will not do so. Moreover, the three theorists which I examine in this essay contend that a moral judgement expresses only a belief; hence, this reconstruction better illustrates the problem which confronts them. I would like to thank Richmond Campbell for bringing this problem to my attention.

[29] Smith 1994,12. For a different, though related formulation of this problem, see McNaughton 1988, 23.

I will use this as a segue into the solution which I will spend the rest of this essay analyzing and criticizing.

As explained above, the view which states that there are no moral truths and that moral judgements do not express beliefs is called non-cognitivism.[30] According to Brink, non-cognitivism is the view that "moral claims do not really make assertions of fact but, rather, express the moral agent's or appraiser's attitudes. Consequently, moral claims can be neither true nor false, there can be no moral facts or true moral claims, and moral knowledge it not possible."[31] Non-cognitivism, then, denies the proposition expressed by 1.

So, what function is served by moral judgements? Under the rubric of non-cognitivism there are a number of different theories regarding the function of moral language and moral judgements.[32] But these theories share some basic similarities. One of the tenets of non-cognitivism is that when we are employing moral language we are not imputing moral properties to agents or acts. Ayer makes the point this way: "The presence of an ethical symbol in a proposition adds nothing to its factual content. Thus if I say to someone, 'You acted wrongly in stealing that money,' I am not stating anything more than if I had simply said, 'You stole that money.' In adding that this action is wrong I am not making any further statement about it."[33] Non-cognitivism of this type holds that to make a moral judgement is just to react emotionally to what one witnesses. A moral condemnation of, say, racist behaviour is to be thought of as an *affective* response to what

---

[30] Ayer 1936, Hare 1981 and Stevenson 1937.
[31] Brink 1989, 3.
[32] For differing versions of the ultimate function of ethical utterances see the works cited in note 29 above.
[33] Ayer 1936, 107.

one believes to be the non-moral facts of the situation. If one takes it that what one has just witnessed is wrong, then one has a negative emotional reaction to it, and one's moral utterance - "That is wrong!" - is an expression of one's emotion.

For this version of non-cognitivism, then, moral experience has two distinct facets: first, we have a set of beliefs about the non-moral facts of the circumstance in which the moral judgement takes place. To be true, these beliefs must fit the world. If they do not, then they are false; and this may undermine the moral judgement we make on the basis of the non-moral facts. Second, we react either negatively or positively to the non-moral facts. The moral judgement is not a further belief about the circumstance, but the result of our emotions; it says something about the appraiser and not about the world.

Non-cognitivists in general are united in their rejection of the proposition expressed in 1. That is, they deny that moral judgments express beliefs and that there is any moral truth. The main virtue of non-cognitivism is that it can explain quite easily the connection between the having of a moral judgement and being motivated to act in accord with it. According to the non-cognitivists, since the uttering of moral judgement is the expression of an emotion or desire, and since the making of a moral judgement is the having of a desire to act in some way, Φ, the maker of the moral judgement, desires to do Φ. And since desires are what motivate the moral appraiser to act, then the appraiser will be motivated to do Φ. Thus, non-cognitivists embrace the practicality of moral judgement, i.e., internalism, and Hume's dictum.

Another way to avoid the moral problem is to deny the proposition expressed by 2 and embrace what has come to be called externalism. The position in 2 is known as

internalism, the view that there is an internal or conceptual connection between having a moral judgement and being motivated to act in accord with it. Externalism denies this. Embracing externalism, and so denying internalism, is the solution adopted by David Brink.[34] According to Brink, externalism is the denial of internalism; "externalism claims that the motivational force and rationality of moral considerations depends on factors external to the moral considerations themselves."[35] Believing that one has a moral obligation will not be sufficient to motivate one to act; one must also desire to do what morality tells one to do before one will be motivated to do what morality enjoins. It is in this sense that motivation is external to moral considerations themselves. Since externalists like Brink deny internalism, they deny 2. Thus, they can embrace moral realism, as Brink does, and also the Humean theory of action explanation.

A third and final way in which to attempt to solve the moral problem is to go after the belief-desire theory of moral motivation, i.e., 3. Proponents of this view are called cognitivist internalists. Cognitivist internalism is the view which accepts the propositions expressed in 1 and 2 while denying 3. This approach, rather than the solutions sketched above, is the solution to the moral problem which I will take up in the remainder of this essay. Those who reject Hume's dictum, at least in part, hold in one form or another that moral judgements are objective and practical. The versions of this position which I consider in this essay are those of David McNaughton[36], Jonathan Dancy[37] and Michael Smith.[38] McNaughton argues, contrary to Hume's dictum, that the mere recognition of a

---

[34] Brink 1989, 37-80.
[35] Brink 1989, 42.
[36] See note 28 above.
[37] Dancy 1993.
[38] Smith 1994.

moral obligation can motivate one to act; and so a belief can motivate. Dancy argues that, though both a desire and a belief must be *present* in order that one be motivated to act in accordance with a moral obligation, a desire is not the *cause* of the motivation. Finally, Smith puts forward an anti-Humean account of rationality; according to this account, some desires are in fact rationally criticizable. Smith's solution to the moral problem turns on his reinterpretation of 2. I will take up these positions in the next three chapters, arguing that each position is flawed in several respects. Hence, I will establish a *prima facie* contention in favour of the view that moral judgements cannot be simultaneously objective and practical.

1.5 Conclusion

Thus far I have sketched perhaps the most pressing problem which faces metaethicists today: the moral problem. In the rest of this essay I will assume, for the sake of argument, that moral realism is true: that there are correct answers to moral questions, that they are made correct by the existence of moral facts and that these facts can be discovered through earnest moralizing. Of course, this view is not without its critics.[39] However, it is not my intention to examine this view. Rather, I wish to take up the problem which besets the internalist moral realist: How can moral judgements be simultaneously objective and practical? I shall pay particular attention to the way in which the three philosophers I examine deal with Hume's dictum. I now turn to this task.

---

[39] See Mackie 1977 and Harman 1977.

CHAPTER TWO

David McNaughton: Internalist Moral Realism[40]

In this chapter, I will recount David McNaughton's arguments in support of a cognitivist internalist thesis. I hope by the end of this discussion to have provided reasons for rejecting nearly every aspect of this theory and for maintaining that Hume's dictum survives McNaughton's challenge.

2.1 Categorical and Hypothetical Imperatives

McNaughton maintains that moral judgements can be true, that moral knowledge is possible and that moral opinions are beliefs about the way the world is, morally speaking. Furthermore, he holds that the mere recognition that one is under a moral obligation to act provides one with motivation or reasons to act, i.e., he holds to internalism.[41] McNaughton thus holds a cognitivist account of motivation. By virtue of the fact that he holds that beliefs - cognitive states - can motivate, McNaughton rejects the notion that an agent is motivated only in the presence of both a belief and a desire. For McNaughton, "the belief [of an agent] that he is morally required to act is sufficient to move the agent to act, without assistance from a different kind of state, a desire."[42]

McNaughton argues that this anti-Humean contention has its basis in common-sense moral thinking: we typically hold that "moral requirements have a claim to our compliance which is independent of what we happen to desire."[43] Moral judgements

---

[40] For the purposes of this essay, I will understand cognitivist internalism and realist internalism to be asserting roughly the same claims. The reader can see these phrases as inter-changeable and synonymous.
[41] McNaughton, 1988, 39-65.
[42] McNaughton, 1988, 49.
[43] McNaughton, 1988, 14-15; see also 48 & 114.

seem to constrain our behaviour. Indeed, sometimes the demands of morality conflict with what we otherwise might desire or prefer to do.

Of course, this way of viewing moral judgements is denied by the externalists.[44] This follows from the fact that they embrace a Humean account of motivation. On this theory, a moral judgement, though it may express a belief, will not be sufficient to motivate an agent to act. A moral judgement motivates an agent to act just if that agent desires to do what morality enjoins; absent the relevant desire, the agent will remain unmoved by moral considerations. Thus, the debate between the externalist and the realist internalist is over what counts as a reason for action: is it a belief alone or the combination of a belief and a desire?

Alternatively, the debate between McNaughton and the Humean can be broadly conceived as one over what type of imperative provides an agent with a reason to perform an action. Imperatives typically take one of the following two forms: 'If you want x, do y' or 'You ought to do y'. Correspondingly, Kant argues that there are two distinct types of imperatives: hypothetical imperatives and categorical imperatives.[45] An imperative is hypothetical if it is based on the putative fact that y is, in the present circumstances, the best means available to bring about the target conditions of the desire for x. The reason for doing y rests on the fact that the doing of y would cause the target conditions of the desire for x to obtain; the oughtness of the imperative is contingent upon the desire. The agent can resist doing the action prescribed by the imperative simply by saying: 'I do not care about x.' However, 'You ought to do y' will be a categorical imperative if you ought

---

[44] See Brink 1989, 37-80, and also Foot 1972.
[45] Kant, 1959

to do y whether or not you have a desire that the doing of y satisfies; that you ought to do y is thus not contingent upon your having any desires. Thus, the agent's reasons for doing y are not extinguished when the agent says: 'I do not care about y.'

An example will help to capture Kant's distinction. Imagine that I have promised to take my sick mother to the pharmacist so that she can collect her much needed prescription, but that when the time comes to act on my promise I do not feel like doing so. At that time, suppose that my sister says to me: 'You ought to fulfill your promise.' Whether the imperative uttered by my sister is hypothetical or categorical depends on the implied reason for keeping the promise. If the reason I am supposed to keep the promise involves an unstated conditional, as in 'You ought to keep your promise, if you want mother to trust you again,' then the imperative is hypothetical; if not, then it is categorical.

In sum, categorical imperatives give reasons for action which are unconditional in the sense of not being contingent upon any present desire of the agent. Categorical imperatives are binding on agents unconditionally, that is, whether or not the agent has a desire to act on the imperative. Hypothetical imperatives, by contrast, are requirements or reasons which are binding on agents only conditionally on the agents having certain desires. Kant held that moral judgements express categorical imperatives, and that categorical imperatives provide an agent with reasons for action.

As suggested above, McNaughton agrees that categorical imperatives provide reasons for action. For he thinks that moral judgements "appear" to provide an agent with

reasons for action independently of an agent's desires.[46] What this commits McNaughton to is the claim that it "appears" that an agent's beliefs about what she is morally required to do are sufficient to motivate her to act, despite what she desires. McNaughton suggests that what justifies this claim is that it coheres well with our moral phenomenology. But this claim appears to beg the question against the Humean! The Humean denies that moral judgements, if they express beliefs about what we ought morally to do, are sufficient to motivate an agent to act. Thus, McNaughton's expostulation begs the question against Hume because it implies that the very concept of a cognitive moral judgement is that it provides an agent with motivation to act. Moreover, since our debate over what counts as a reason for action just is a debate over whether a moral belief alone can provide one with a motive to act, to assert that cognitive moral judgements provide an agent with a motive is to dogmatically assume the very claim over which Hume and McNaughton disagree. But this charge would be too quick, since McNaughton holds that it only "appears" that this is so. He can, and, in fact does, attempt to demonstrate that cognitive moral judgements[47] provide reason for action by way of refuting Hume's dictum. I detail this argument in the next section.

There is one difficulty with McNaughton's contention that categorical imperatives provide a reason for acting, namely, that if one is motivated by a belief that one has a moral obligation - and a desire does not figure into the reason for acting upon one's moral obligation - then we might plausibly conclude that the person acts according to the

---

[46] McNaughton 1988, 48.

[47] Hereafter, I drop the term cognitive from the phrase 'cognitive moral judgements'. When I use the phrase moral judgements, I mean cognitive moral judgements.

dictates of morality, but does not desire to do so.[48] McNaughton sees this consequence of his position as "excessively austere and surely implausible."[49]

In response, he argues that the person who acts morally - the virtuous person - can be seen as doing what morality enjoins, not as a soldier who begrudgingly does what he is told because he is told to do it, but as a person who takes joy and pleasure in doing what morality enjoins.[50] Moreover, the virtuous person can even be said to have desired to do what he did. McNaughton suggests that "to ascribe such a desire to the agent, after he acted, is merely to acknowledge that his moral belief was here sufficient to motivate him."[51] It seems, then, that McNaughton has made a concession to the Humean, since he proposes that for an agent to be suitably motivated, both a belief and a desire must be present.

Does this not concede too much to Hume, and so undermine the contention that moral judgements express categorical imperatives, and that categorical imperatives are sufficient to motivate an agent to act? McNaughton argues not. Though we ascribe a desire to an agent when she acts intentionally, it does not follow that motivating reasons for action are constituted by beliefs and desires. McNaughton agrees with Hume that all intentional actions take place in the presence of a belief and a desire, but he denies that a desire is always a Humean desire. The account of desire that McNaughton espouses does not see a desire as a distinct existence in the way Hume thought, nor as a source of

---

[48] McNaughton 1988, 50 & 106.
[49] McNaughton 1988, 50.
[50] McNaughton 1988, 50 &106.
[51] McNaughton 1988, 50.

motivation, but merely as something that is ascribed to an agent when she has acted intentionally.[52]

This concession to the Humean is paradoxical, though. McNaughton contends that moral judgements give agents reasons for acting quite independently of what they desire or prefer. At the same time, however, McNaughton, suggests that a person who acts morally - the virtuous person - in fact *has a desire* to act morally. And, there is a tension here: on the one hand, moral judgements give agents reasons for acting independently of what they desire; on the other hand, given McNaughton's account of desire, agents act on their moral judgements because they desire to do so. What, then, is McNaughton up to?

I do not think that this tension is really a problem for McNaughton since he does not assume that if one desires to act in accord with a categorical imperative, it is no longer a categorical imperative. To put it another way, categorical imperatives give an agent reasons for action independent of what they may otherwise desire or prefer to do, but an agent may also desire to do what the categorical imperative enjoins, and it does not follow from the fact that because an agent desired to act in the way prescribed by a categorical imperative that the imperative is no longer categorical. However, two things follow from holding that a moral judgement is a categorical imperative and that a categorical imperative provides reason for action. First, a desire need not figure into the reasons that an agent acted on a categorical imperative, even if the desire to do what a categorical imperative enjoins is present in the agent. Second, and perhaps this follows from the last statement, an agent should act in accord with a categorical imperative even

---

[52] McNaughton 1988, 50 &106.

if she does not desire to do so and even if she desires to act in a way contrary to the imperative.

But here is where McNaughton's account of desire becomes problematic. If he holds that categorical imperatives provide reasons for acting independent of an agent's desires, then he will have to hold that an agent should act in the way enjoined even if he desires otherwise. To illustrate how this is a problem for McNaughton, I will employ an example. Imagine that I have promised to visit my sick Grandmother in the hospital. Since I am morally obliged to fulfill that promise, I am under a categorical imperative. Yet, suppose that when it comes time act on the promise, I no longer desire to do so. Rather, I want to attend the football game to which I have tickets, and I have had a burning desire to do so for a long time now. Now, I see myself with a moral obligation, and my moral obligation, since it expresses a categorical imperative, provides me with reasons for acting despite what my desires may be. In this case, I desire not to act in the way prescribed by the categorical imperative. I am torn; I do not know what to do. However, after much deliberation, I decide that I will act in accord with my obligation despite the fact that I have no desire to act in accord with it. Now, on McNaughton's account of desire, since I acted in accord with what my moral judgement enjoined, I desired to do what I did, because for McNaughton "to ascribe a desire to an agent, after he acted [in accord with a categorical imperative], is merely to acknowledge that his moral belief was here sufficient to motivate him."[53] Yet, in this case, my moral belief did in fact motivate me to act, but I did not desire to do what my moral belief enjoined. But,

[53] McNaughton 1988, 50.

since McNaughton says that if a moral belief was sufficient to motivate me to act, then I desired to do what I did, I desired to act the way I did. Thus, on McNaughton's account, I desired to act morally. But according to the story I just told I *did not* desire to act morally; I desired to go to the football game. McNaughton's account of desire ascribes a desire to an agent when no such desire is present in the agent when he acted in accord with a categorical imperative. If this is a consequence of McNaughton's theory, then it must be flawed. And, if it is flawed in this way, then it must be false; if it is false, then it must be jettisoned.

Of course, the Humean might quarrel with the example I have constructed above because it seems psychologically unusual, given that on Hume's theory every action issues from a motivating reason which is constituted by both beliefs and desires. I agree that this is so, but despite the fact that it seems psychologically unusual, McNaughton is still committed to the fact that such an example is at least a possibility: he thinks that a moral belief by itself can motivate without the help of a desire and that one ought to be motivated by a moral judgement whatever one desires.

Notice that McNaughton cannot simply jettison his account of desire because, if he does, then he falls prey to the objection that provided the impetus for his account of desire in the first place. That objection, remember, was that if one holds that a categorical imperative provides an agent with a reason for action, then one is committed to the view that "the virtuous person has no desire to act in the way he does, when he acts in virtue of his belief that a certain course of action is morally required."[54] Thus, if McNaughton

---

[54] McNaughton 1988, 50.

rejects his account of desire, then he is still susceptible to this objection to his account of moral judgements and, hence, is left with an account of morality that is, in his words, "excessively austere and surely implausible."[55]

McNaughton might suggest in response to my objection that since I did act in accord with my promise in the above example, I must have desired to do so. By saying this, he could argue that the example I gave is impossible: no one is capable of acting in a way that is contrary to their desires. There are two rebuttals to this objection. First, if he concedes that one cannot act contrary to one's desires, then he cannot hold that moral beliefs by themselves give us reasons for action.[56] That is, McNaughton cannot hold that categorical imperatives give an agent reasons for acting independent of her desires. Furthermore, if he concedes that action always requires a desire, then he is conceding that a desire has some sort of motivational "oomph" or comprises part of the reasons for why agents act. And, if he concedes this, then he must concede that it is not the moral belief by itself that is motivating agents to act. But he wants to hold that the moral belief is what motivates, so he cannot use this as a response to my objection. Therefore, McNaughton's account of desire must be false and should be rejected.

It seems that my example forces McNaughton into a dilemma. Either he must admit that it is possible for one to act on moral judgements - since they give reasons for action - despite the fact that one does not desire to do so, or desires to do something else. Or, he must concede that one acts on a moral belief only if one desires to do as the belief

---

[55] McNaughton 1988, 50.

[56] What about a case where an agent has no desires on the matter either way? If this is the case, then the agent cannot act, since she has reasons for acting both ways, or no way at all. But if this is the case, then she will not act either way, because she has no determinate reason for doing so. Hence, it still follows that a categorical imperative alone will not be sufficient to motivate an agent to act.

tells one. If McNaughton takes the first horn, then he is saddled with an account of moral judgements which is "excessively austere and surely implausible."[57] If he takes the second horn, then he is committed to saying that moral beliefs are not sufficient to motivate an agent to act, and that moral judgements do not provide an agent with reasons to act despite the agent's desires. Either way, it seems that he is left with the prospect of abandoning something central to his account of morality.

But perhaps McNaughton might have a reply to the second horn of the dilemma. He might argue that I am motivated to Φ and, hence, desire to Φ because I believe that Φ-ing is my moral obligation. My motivation to Φ does not exist in virtue of the fact that I have an antecedent desire to Φ; my desire to Φ is caused by what my moral obligation enjoins. My desire to Φ thus explains my belief that I ought to Φ. However, there are two objections to this response. First, it is not at all clear how it is that a belief can cause a desire since beliefs and desires are independent existences. Moreover, McNaughton would have to admit that if a belief causes an agent to have a desire, then whenever an agent has that belief, then he has that desire. But this seems implausibly strong: there are cases in which an agent can hold a belief without having the concomitant desire, e.g., when an agent is weak-willed. McNaughton would need to explain how it is that on some occasions a belief causes a desire, while on other occasions it does not. It seems that no explanation is forthcoming. Second, even if we were to concede that a belief is capable of causing a desire, this contention undermines McNaughton's contention that a moral belief alone provides an agent with reasons for action and, hence, the claim that a categorical

[57] McNaughton 1988, 50.

imperative provides an agent with reason for action. This follows from the fact that if a belief causes a desire and a desire is what moves an agent to act, it is not the moral belief, but the desire, which moves the agent to act.[58] Thus, McNaughton's account of desire seems defective.

2.2 A Cognitive Account of Motivation

McNaughton thinks that he can offer us a cognitive account of motivation. He thereby denies that a cognitive state, i.e., a belief, cannot motivate one to act. On this view, a desire does not need to be added to the explanation of why an agent acted the way she did; nor is a desire required in order that an agent be motivated to act. Thus, McNaughton rejects the Humean account of action explanation and motivation which states that an agent is motivated just if she has both a desire and a means-end belief.

However, McNaughton must meet one objection before he can vindicate his cognitive account of motivation. In the last chapter (1.2), I explained, following Mark Platts, that beliefs and desires are radically different kinds of states, since they have different and opposing directions of fit. According to this explanation, beliefs are responsible for representing the world truly or falsely, while desires have no such responsibility. Desires, rather, are directed at changing the world. They do not represent the world truly or falsely, though they might be said to represent the world as it would be if the target conditions of that desire obtained. Beliefs aim at the truth, while desires aim at their own satisfaction. If a belief fits the world, then it is true; if the target conditions of

[58] For more on this argument see section 2.3.

a desire obtain, then the desire is satisfied. Beliefs, then, must fit the world to be true while the world must be changed to fit a desire, if a desire is to be satisfied.

McNaughton argues that this account of the distinction between beliefs and desires poses a problem for the cognitivist account of motivation. We have characterized beliefs and desires in such a way that we can only explain an agent's action if we ascribe to her some end or goal. And, this can be done only insofar as we can ascribe a state to her whose satisfaction would consist in the world being altered to fit it. Desires are states with this direction of fit. Thus, if we want to explain why an agent acted, we must hold that she desired to do the action, or desires an outcome of doing it.

However, McNaughton's cognitivist account of motivation purports to explain an agent's action by ascribing to her a belief about what she is morally required to do. But a belief has a certain direction of fit - that the state must fit the world. And this is the wrong direction of fit to explain an agent's action. In order to explain why an agent acted, we must ascribe to her a state with the direction of fit of a desire, namely, a state which is such that the world must fit it. Ascribing to the agent a belief, therefore, fails to explain why the agent acted to bring about a change in the world; only the ascription of a desire can do this since only it has the correct direction of fit.

In reply to this objection, McNaughton argues that "some cognitive states can be seen as having. . .directions of fit running both ways."[59] This sort of state would be such that the world must fit with it, insofar as it is desire-like, and such that it must fit the world insofar as it is belief-like. McNaughton maintains that if an agent is aware

[59] McNaughton 1988, 109.

(believes) that he is under a moral obligation to act in a certain way, then he *sees* the situation as demanding a response; if he is disposed to act, then he must be in a state with the direction of fit of a desire, namely, one to which the world must fit. This state will only be satisfied if the agent acts to alter the world to fit the state. However, McNaughton also wants to claim that the agent's conception of the situation is "purely cognitive."[60] This follows from the fact that the agent has a belief that he is morally obliged to act and so his state must fit the world. The belief that he is under a moral obligation will be true just if it represents accurately the way the world is, morally speaking. McNaughton thus commits himself to the claim that "awareness of a moral requirement is a state which must be thought of as having directions of fit facing both ways. The agent's conception of the situation reveals to him both that the world is a certain way and that he must change it."[61]

One technical term for a state with two directions of fit is a "besire."[62] However, against prevailing philosophical orthodoxy, McNaughton does not call such a state a 'besire', but a belief. But on what grounds is he entitled to do this? McNaughton suggests that though it is a sufficient condition of a state's being a belief that it has the direction of fit that the state must fit the world, "it does not lose the status of a belief if it happens to have another direction of fit as well."[63] But can we not say the same of a desire? A sufficient condition of a state's being a desire is that it has the direction of fit that the world must fit it. But we might also say that though it is a sufficient condition of a state's

---

[60] McNaughton 1988, 109.
[61] McNaughton 1988, 109.
[62] The term "besire" comes from James Altham. See Altham 1986. Michael Smith calls a state with two directions of fit a "quasi-belief". See Smith 1987: 56.
[63] McNaughton 1988, 110.

being a desire that it has the direction of fit that the world must fit it, it does not lose the status of a desire if it happens to have another direction of fit as well. Hence, following McNaughton's own reasoning, surely a state can be called a desire despite the fact that it has both directions fit. Perhaps McNaughton could embrace such a conclusion, but it would severely cost him, for he wants to hold that a state with two directions of fit is in fact a belief. He wants to hold this because he wants to hold that a moral belief is sufficient to motivate or give reasons for an agent to act. Therefore, McNaughton cannot make this move. But then he does not have an independent argument as to why such a state should be called a belief rather than a desire or a besire. Without an independent argument for calling such a state a belief rather than a desire such a state simply cannot rightly be called a mere belief.

But perhaps McNaughton can claim that the state with two directions of fit can be called a belief since he argues that an agent's conception of the situation is "purely cognitive." As McNaughton puts it: "beliefs...are cognitive states - they are representations of the way we take the world to be."[64] However, in saying that a state with two directions of fit is purely cognitive like a belief does he means that it only serves to represent and not to motivate? Surely not. But, then, he fails to provide us with an argument to the effect that something can be both solely representational and simultaneously motivational, since for something to be purely representational is for it to be representational but not motivational. And, he needs an argument to this effect in order to demonstrate that a moral judgement, which expresses a belief about the way the world

[64] McNaughton 1988, 21.

is morally, is sufficient to motivate an agent to act. Although I think these objections tell against McNaughton's contention that a state can have both directions of fit, I think that the notion suffers from more intractable problems; and that the only way in which McNaughton can escape these problems is a way which undermines his claim that moral beliefs are sufficient to motivate an agent to act.

2.3 The Cognitive Account of Motivation Undermined

Michael Smith argues that the notion of a state with both directions of fit is incoherent.[65] Smith justifies this conclusion by examining the differences between beliefs and desires in terms of direction of fit. Smith argues that the difference "comes down to a difference between the counterfactual dependence of a belief and a desire that *p*, on a perception that *not-p*: roughly, a belief that *p* is a state that tends to go out of existence in the presence of a perception that *not-p*, whereas a desire that *p* is a state that tends to endure, disposing the subject in that state to bring it about that *p*."[66] With this in mind, consider the following example. Imagine that I desire that x obtain. If I see that not-x obtains, then my desire that x obtain will tend to persist, and, hence, motivate me to act to bring about the target conditions of my desire, namely, x. This follows from the fact that the world does not yet fit my state. However, if I see that x obtains, i.e., if I see that the world fits my state, then since the target conditions of my desire already obtain, my desire will extinguish.

How does this make the notion of a state with both directions of fit incoherent? For one to have a desire that x obtain, one must necessarily believe that not-x obtains in

---

[65] Smith 1994, 116-125. See also Michael Smith 1987, 55-57.
[66] Smith 1987, 54.

the world; for if it did not, if instead x obtained, and, if I saw this, the desire should be extinguished. But if a state has both directions of fit, the target conditions of the state would have to both obtain and not obtain. The target conditions of my state would have to obtain in order for the state to fit the world, and it would have to not obtain, i.e., not fit that world, for me to be motivated to bring about its target conditions. Moreover, as Smith puts it: "A state with both directions of fit would . . . have to be such that *both*, in the presence of such a perception, it tends to go out of existence, *and*, in the presence of such a perception, it tends to endure, disposing the subject that has it to bring it about that [x]."[67] Thus, a state with both direction of fit requires that the same state of the world would have to obtain and not obtain and the same mental state endure and not endure, in order to motivate an agent to bring its target conditions about. But this is absurd, for something cannot both be x and not-x simultaneously. Thus, the contention that there is a state with both directions of fit is contradictory, and hence incoherent. It is important to point out here that a state with two directions of fit is incoherent for two reasons. First, because it requires that the same state persist and not persist in the presence of the same relevant perception. Second, because it requires that the world be a certain way and that it not be a certain way.[68]

How might McNaughton reply to these objections? I think a plausible reply can be mustered by looking at a concrete example. Imagine that I see a young child struggling in the water. If I do not throw her a life-preserver, she will surely perish. Suppose that it is also true that I ought to save the drowning girl because morality requires this of me,

---

[67] Smith, 1987, 56.

[68] Smith does not explicitly assert this second proposition, but I think that it is implicit in his argument.

and I believe that I ought to save the drowning girl. McNaughton argues that my conception of the situation is enough to move me to act; my mere belief that I have a moral obligation provides me with the motivation to act. In this case, since I believe truly that I ought to save the drowning girl, my state fits the world. Of course, if this state has two directions of fit, then it will also have to be the case that my state does not fit the world so as to motivate me to act. The part of my state which motivates me to act is the part that represents it that *that the girl be saved*; this part of my state motivates me to bring it about that the world fits the state, and it does so because it does not yet fit the world. Once I act to bring about the target conditions of this part of the state, the world will be such that it fits the state. Further, I am not motivated to bring it about that *It ought to be that the girl be saved.* For this proposition is represented in the belief-like part of my state.

The only way such a state would be contradictory would be if the state motivated me to change the world so as to make the belief-like part of my state or the moral judgment true, namely, '*It ought to be that the drowning girl be saved.*' But this is not what I am acting to bring about; I am acting to make true the state of affairs *prescribed* by the belief part of my state, namely, *that the drowning girl be saved.* If it was the case that I was acting to make the belief-like part of my state true, then both the belief-like part of my state would have to be true *and* not true simultaneously, and the state would have to endure and not endure, if I was to be motivated. However, in the example given above, I am not acting to make the belief-like part of my state true. Rather, I am acting to make the state of affairs prescribed by the belief-like part of my state true. Since, the target

conditions prescribed by the belief part of my state do not yet obtain when the belief-part is true, the state still endures and so motivates me to act. Thus, a state with both directions of fit is not contradictory and so not incoherent.

I think this move gets McNaughton off the hook regarding the incoherence of states with both directions of fit. But it does so in a way which undermines McNaughton's overall contention, namely, that cognitive moral judgements are sufficient to motivate action. The reply commits him to the assumption that a moral judgement would have two intentional objects. The mental state that it is wrong not to save the girl has the propositional content that it is wrong not to save the drowning girl and a prescriptive content, namely, that the drowning girl be saved. Since the propositional content of the state and the recommendation of the state differ, and I am motivated to bring about the latter, the objection that a state with two directions of fit is contradictory is avoided. However, it avoids the objection at the cost that it is not the propositional content of the state alone that serves to motivate the agent to act; hence, a belief will not be the sole source of motivation. Since the propositional content of the state - that it is wrong not to save the drowning girl - represents the moral judgement, that it is wrong not the save the drowning girl, the reply has the consequence that it is not just what the moral judgement is about that serves to motivate the agent to act. What motivates me to act is what the propositional content of my state recommends, and not the propositional content alone. Yet, McNaughton wants to maintain that it is the moral judgement by itself that serves to motivate the agent to act. According to McNaughton, what motivates me to save the drowning girl is that it is right to save the drowning girl. But in order to avoid the

prima facie contradictory nature of a state with two directions of fit, he must concede that the propositional content of my state, namely, 'It ought to be that the drowning girl be saved', does not motivate. The states of affairs prescribed by the propositional content of the state motivated me to act, and not the propositional content it self, and so not the moral judgement. Hence, the belief that I am under a moral obligation by itself, at least on McNaughton's account, does not serve to motivate. Moreover, if the state can be divided up in the way that I have just suggested, then we can see that one part is a belief and the other is a desire, and this is just what the Humean wants. Thus, his cognitivist account of motivation is undermined. Therefore, McNaughton's cognitivist account of motivation does not provide a threat to the Humean account of motivation.

2.4 The Case of the Masochist

Thus far I have argued that McNaughton's cognitivist account of motivation is flawed. But McNaughton has one more attack on Hume's dictum, namely, that it is explanatorily impotent.[69] If he can prove this, then he might be able to at least show how a cognitivist account of motivation is more plausible than Hume's account at least in terms of explaining action. This will not, of course, shield McNaughton from my attack on his account of motivation. It will merely show that Hume's account of action explanation is no better off than his own.

Hume, remember, argues that there are two distinct mental states: beliefs and desires. Beliefs tell us how the world is at present, and how it might be changed to make it fit with the target conditions of our desires. Hume employs this account of human

[69] McNaughton 1988, 110-113.

psychology to explain intentional human action. For Hume, an action is the product of both beliefs and desires. Desires represent the way the world is to be when their target conditions obtain, and a belief represents how the world is and how it has to be changed to make it fit with our desires. Thus if one is to be motivated to act one must have both a means-end belief and a desire. McNaughton calls this Hume's belief-desire theory of action explanation.

McNaughton argues that on Hume's belief-desire theory of action explanation, it is logically possible for an agent to desire any state of affairs. McNaughton argues that "if an agent's beliefs about a desired state of affairs or his way of conceiving of what he desires is quite independent of his desiring then he can consistently desire anything, no matter how he conceives of it. There is no limit to the combinations of beliefs and desires that are possible."[70] McNaughton suggests that this contention may be suggested by empirical evidence: people do desire all kinds of strange things. Take, for instance, masochism. The masochist is a person who claims to desire to be in pain, which is something most of us try our utmost to avoid. On the Humean account of action explanation where desires are non-cognitive states, this desire seems perfectly intelligible. The fact that some people desire pain while most avoid it is no more strange than the fact that some people like steak while others abhor it.

But McNaughton argues that the desires which the Humean says are intelligible in fact appear to be unintelligible; we find them bizarre. For instance, we cannot understand why a masochist desires being beaten and whipped. Yet, according to "the

---

[70] McNaughton 1988, 111.

belief-desire theory of action explanation, we are wrong to be puzzled."[71] For in this theory, it is enough to explain why the masochist acted to state that he desired to be in pain. McNaughton argues that the fact that the belief-desire theory can explain this action so easily must mean there is something wrong with this theory, for we are in fact puzzled by such desires. And, "if we cannot make sense of that desire then we cannot make sense of an action explanation which appeals to that desire."[72] Though the Humean account of desire can explain what the agent is aiming to bring about, it does not explain his action fully. McNaughton claims that this follows from the fact that "we have no insight into why being in the desired state should bring satisfaction."[73]

McNaughton alleges that we require a more satisfactory account of explanation of action than the belief-desire theory allows us, and that the cognitive account of motivation can provide this. On the cognitive account of action explanation, we explain the masochist's action by appreciating the masochist's view of whipping, and we do this by discovering why the masochist finds whipping attractive. There are several paths to this goal. For example, we may point out that "some things that are normally experienced as mildly painful can be enjoyable when we are sexually excited, or that punishment can have the pleasurable effect of relieving guilt."[74] McNaughton suggests that we come to understand the desire for pain by coming to appreciate, though not sharing, the experience of the masochist whose actions we are trying to explain. The agent's conception of the situation makes the desire intelligible. Once we have grasped the

---

[71] McNaughton 1988, 112.
[72] McNaughton 1988, 111.
[73] McNaughton 1988, 111.
[74] McNaughton 1988, 112.

attractiveness of the desire, we come to understand why that person was motivated to act to bring about the target conditions of his desire. McNaughton contends that we do not therefore, contra Hume, need to add anything to his conception to explain his action. We do not need to ascribe to the agent a non-cognitive state. If we were to entirely share the masochist's conception of the situation, then we would be motivated to act as he does. Thus, McNaughton concludes, all we need to explain action is the cognitive account of action explanation.

I think that there are two responses to this objection. First, it is not clear that we really need a cognitive account of action explanation to explain how it is that some desires, though they seem *prima facie* unintelligible, are intelligible. Hume could simply deny that the desires that McNaughton claims are unintelligible, e.g., the desire to drink a can of paint, are really unintelligible. For Hume, that we have a given desire is just a fact about us. Furthermore, Hume can quite easily explain a desire and its corresponding action by employing his belief-desire theory of action explanation. If we find that a certain agent thinks that a certain state is desirable, then it seems to be the case that we will come to realize or know that he will be drawn to bring about the target conditions of that desire; and hence, we come to understand why he came to act the way he did. He acted that way because he desired the outcome of that action. Someone's desiring something is evidence that he will act in ways that will satisfy that desire.

However, I might agree with McNaughton's claim that if we are really to understand an agent's desires and why he acts to bring their targets about, then we might need to get inside his head, so that we can really see the desire and the action from his

point of view. In the case of the masochist, we might need to know what the agent believes about being whipped to discover why he desires to be whipped. We do this by, for instance, coming to discover that the masochist believes that being whipped is pleasurable and exciting, or that it relieves some kind of guilt. For the Humean these are the agent's beliefs about the situation, and these beliefs explain why the masochist finds being whipped attractive. The fact that the masochist desires being whipped exists in virtue of the fact that he believes that it will, say, relieve guilt or increase pleasure, and thus explains why he desires being whipped. If the Humean can do this, then he can make the desire intelligible in just the way that McNaughton requires.

Now, we know that the masochist believes that being whipped will relieve his guilt or cause him pleasure, and that explains why he desires being whipped. However, just knowing how the masochist conceives of the situation will not be sufficient to explain why he acts to bring about the target conditions of his desire to be whipped: we must also know that the masochist desires (in the non-cognitivist sense) pleasure or the relief of guilt. Hume can explain the seemingly unintelligible desire to be whipped which the masochist has. He explains it by citing a further belief which the masochist has, namely, that being whipped relieves guilt or causes pleasure. But what is salient about Hume's explanation of the agent's action is the positing of a non-derivative desire to relieve guilt or experience pleasure, which motivated the masochist to desire to be whipped in the first place. Without this further desire, the desire to be whipped does not make sense. The desire to relieve guilt explains the desire to be whipped, and this desire is explained by appeal to the belief that being whipped relieves guilt or causes pleasure.

But the former desire needs no explanation. So it will not be enough to explain why the agent acted by citing his beliefs about the situation; what we need is a non-derivative desire. Thus, even on the belief-desire theory of action explanation we can make sense of the masochist's desire to be whipped, and thus explain why he acted to bring about that he be whipped. We need to know therefore, not only how an agent conceives of things, e.g., his belief that whipping relieves guilt, but also that he desires the relief of guilt. Therefore, I think that McNaughton's objection to the Humean account of action explanation fails.

2.5 Conclusion

In this chapter I have argued that David McNaughton's cognitivist internalist account of motivation is flawed because it relies on the plausibility of the notion of a state with two directions of fit, which itself is flawed. Furthermore, the only reply that he can give to avoid the flaws with a state with two directions of fit undermines his contention that a moral belief is sufficient to motivate an agent to act. I have also argued that his attack on the Humean account of action explanation can be defused by the Humean. Thus, we should reject McNaughton's argument for internalist realism and his attacks on Hume's dictum. At the end of the last chapter and the beginning of this chapter I argued that if McNaughton's internalist realism is to move beyond its mere intuitive plausibility, it must provide us with sound philosophical arguments. If what I have said about his thesis is correct, then he has failed at this task, and so we should reject his account of internalist moral realism. However, it would be too hasty to say that internalist realism has gone down in an ignominious defeat. In the next chapter I examine Jonathan Dancy's attempt

to provide us with an argument for internalist realism. Dancy's position does not employ the same arguments as McNaughton, and so must be treated as an altogether new attempt to vindicate internalist realism.

CHAPTER THREE

Jonathan Dancy: The Pure Theory

In the last chapter, I argued that McNaughton failed to establish the truth of cognitivist internalism. In this chapter, I will take up Jonathan Dancy's attempt to provide arguments in favour of this same thesis. Dancy suggests that there is a strong intuitive argument in favour of cognitive internalism. He considers this argument to be 'strong', because there is "a sense [in which] morality is essentially practical, so that it would be odd for someone to say 'This action is wrong but I do not see that as at all relevant to my choice.'"[75] This intuition provides the point of departure for Dancy's defense of cognitive internalism. Of course, Dancy does not stop here; he attempts to demonstrate that there are good philosophical reasons for embracing internalism. In this chapter, I present those reasons. I reject Dancy's internalism and, in the process, give at least *prima facie* reasons in favour of externalism.

3.1 The Pure Theory

Dancy purports to provide us with a version of cognitive internalism. As should be familiar by now, this view has two features: first, the claim that moral judgements express beliefs; second, the claim that moral judgements necessarily motivate or provide reasons for action. What should also be familiar by now is that this thesis clashes directly with Hume's dictum. Hume maintains that every completely motivating state is a combination of a belief and a desire. Hume also suggests that beliefs are motivationally

[75] Dancy 1993, 4.

impotent, whereas desires are essentially motivating. Thus, beliefs and desires conjoined constitute a reason for action.

Dancy agrees in part with Hume, maintaining that a completely motivating state has two elements, labeled belief and desire. But he disagrees with Hume over the way in which beliefs and desires are to be conceived. Dancy suggests that the two be conceived in broadly cognitivist terms. On Dancy's view, what is required for an agent to be motivated to act is that there be "two distinct 'representations' in the agent."[76] The first is the agent's belief(s) about the way the world is at present; the second is the agent's belief about the way the world will be once the action is completed. An agent needs to hold the first representation because without it she would not be aware of the circumstances within which she acts, nor whether her action will lead to any change. The second representation is needed because without it an agent would be blind to what she is trying to bring about.

An example will demonstrate just what Dancy has in mind. Imagine that I believe that my girlfriend is in extreme pain, and I also believe that were I to give her the appropriate dose of morphine, her pain would be mitigated. There are two representations here: (1) The belief that my girlfriend is in extreme pain; (2) The belief that were I give her the morphine, her pain would be mitigated. Dancy contends that the two representations are enough to motivate me to act.

Two questions immediately arise with respect to Dancy's claim that the two representations motivate action: (1) What is it about them that motivates the agent to act? (2) Where does the notion of a desire figure into this account? In answer to the first

---

[76] Dancy 1993, 13.

question, Dancy argues that what motivates the agent to act is "the gap between the two representations."[77] In the case above, then, what motivates me to administer the morphine to my girlfriend is the gap between the belief that she is in pain and the belief that were I give her the morphine, then she will be relieved of her pain. Furthermore, for the agent to be motivated there does not need to be any instance of a "recognizable version of Hume's conception of desire."[78]

In answer to the second question, Dancy maintains that "the occurrence of the desire is *the agent's being motivated* by the gap."[79] Dancy's account of desire differs from Hume's in that the desire is "not part of the causal story that takes us from beliefs to action. The desire here is to be conceived as a distinct event, that of the agent's being motivated by certain conceptions. As such it is caused by those conceptions, but though necessary for eventual action it is not a cause of that action; it does not pull its own weight in the causal story."[80]

In essence, this is Dancy's account of motivation. Dancy does not provide a detailed elaboration of his account of motivation. As Dancy explains: "I shall never offer an explicit argument in favour of that theory [namely, his theory of motivation], nor argue against its Humean rival."[81] Instead, his view must be gleaned from the way in which he defends his theory against Humean-type attacks. In the next four sections I take up Dancy's defenses, arguing that they all fail. In the fifth section I attempt to demonstrate that Dancy's account of motivation is fraught with problems.

---

[77] Dancy 1993, 19.
[78] Dancy 1993, 14.
[79] Dancy 1993, 19.
[80] Dancy 1993, 20.
[81] Dancy 1993, 20.

3.2 Amoralism, Wickedness and Lucifer Himself

If Dancy's contention is that moral judgements, which express beliefs, are sufficient to motivate one to act, then it rules out as impossible the figures of the amoralist and the wicked person. The amoralist is one who is capable of distinguishing between right and wrong, but is not concerned with such considerations and, hence, is not motivated by them. The wicked person, like the amoralist, knows the distinction between right and wrong, but sees the wrongness of an action as a reason for doing it, and the rightness of an action as a reason for leaving it undone.[82]

David Brink argues that the mere possibility of such people poses a problem for internalism in the following way: Internalism is the view that it is part of the concept of morality that a belief about what we are obliged to do necessarily provides an agent with a reason or motivation for action. If people such as the amoralist and the wicked person are possible, then internalism is false.[83] The amoralist and the wicked person are possible. Therefore, internalism is false.[84]

Brink argues that externalism can account for the existence of such people because the externalist holds that "the motivational force of moral considerations [is] a matter of contingent psychological fact, depending on the beliefs and desires agents happen to have."[85] Externalism thus takes it as at least a *possibility* that an agent can believe that he is under a moral obligation, and not thereby be motivated to act. *A fortiori*, that such people are possible is an argument in favour of externalism. Dancy seeks to

---

[82] See Dancy 1993, 4-6; see also McNaughton 1988, 134-146.
[83] This formulation isn't quite apposite for the case of the wicked person. The wicked person is motivated by such considerations, but in the wrong way.
[84] Brink 1989, 45-50.
[85] Brink 1989, 49.

establish against this objection that neither the genuine amoralist nor the wicked person, as conceived in the objection, are possible. And, since they are not possible, they do not pose a problem for the internalist.

Regarding the amoralist Dancy says this: "we may all admit that there is the person who sees the institution of morality from the outside, as something whose claims on us he rejects," but this person "*does not* accept the moral judgements whose relevance he denies . . . he merely knows what judgements would be made by others."[86] The amoralist, then, is not a person who believes that what he is doing is wrong; rather, he is one who simply rejects conventional morality. This person is no trouble for the internalist.

Dancy's argument amounts to no more than the dogmatic assertion that the amoralist - one who accepts the dictates of morality but is not motivated by them - is not possible. But this seems a jejune response to Brink's objection, since it begs the question. According to Brink, it is implausible to assume that a person who remains unmoved by moral considerations is merely to be seen as rejecting conventional morality. For as Brink puts it: "We can imagine someone who regards what we take to be the moral demands as moral demands - not simply as conventional morality - and yet remains unmoved."[87] But perhaps Dancy's objection is more subtle. He might be suggesting that if one were *really* to believe that one is under a moral obligation, then one would be moved to act. In other words, the fact that one is not moved by the moral obligation that one supposedly believes one has, proves that one really does not believe that one has a moral obligation at

---

[86] Dancy 1993, 5; emphasis added.
[87] Brink 1993, 48.

all. If one did believe it, then one would be motivated. However, if this is Dancy's argument, then he has merely assumed the truth of internalism and thereby begged the question against the externalist. This response is question-begging, because it states that if one were under a moral obligation then one would thereby be motivated. But externalism denies just this thesis. Externalism claims that it is possible for one to believe that one is under a moral obligation, but not be moved by that consideration. For externalists, something external to morality must be present for an agent to be motivated.

However, perhaps it is the externalist who is really making a dogmatic assertion here to the effect that his conception of the amoralist is possible; hence, perhaps the externalist is begging the question against the internalists' conception of the amoralist. It seems that we are at an impasse. Is there any way out? The externalist can simply turn the matter over to empirical evidence. There do seem to be people in society, e.g., sociopaths, killers and mad people, who believe that what they are doing is wrong, yet regard this belief as irrelevant to their actions.[88] Empirical evidence (i.e., that there are such people) and experience seem to confirm the externalist's conception of the amoralist and, at the same time, tell against Dancy's claim that the amoralist as conceived by Brink's objection could not possibly exist. Dancy thus fails to come to grips in a serious way with the possibility of the amoralist; he merely denies what we have good empirical evidence to assert. However, Dancy might dissent from the externalist's reading of the empirical data, arguing that the person who is thought to be an amoralist is really not the type of

[88] For an interesting account of a real-life sociopath, see Watson (1987). Watson recounts the story of thrill-killer Robert Harris who claims that he knew that what he was doing was wrong and that he chose to do it despite this fact.

amoralist that the externalist needs to undermine internalism. Dancy can simply argue that the way he reads the empirical data concurs with his concept of the amoralist. Thus, we are again at a deadlock. Both Dancy and the externalist need to provide an independent, non-question begging argument regarding the correct account of the amoralist. However, since it seems that neither Dancy nor the externalist can do this, we must search for another way to break the impasse between the two theories. The possibility of breaking the impasse comes through reflection on another kind of person - the wicked person - who undermines internalism. I now turn to that argument.

Dancy offers the following objection to the argument from wickedness. Dancy maintains that although we can conceive of the notion of a "wicked person", it is impossible that such a person could actually exist. That is, there could be no person who takes the badness of his actions as a reason for carrying them out. For example, although we can have a notion of a person, such as Lucifer, who pursues evil for its own sake, it is in fact impossible that Lucifer could exist as conceived. It is simply impossible for an agent to pursue evil for its own sake.[89] Dancy suggests that Lucifer does not "pursue evil . . . for its own sake (*sub specie mali*) . . . but for extraneous reasons, such as the fact that it is his last remaining hope of an empire."[90] According to Dancy, "this renders Satan's pursuit of evil comprehensible, but at the cost of making it useless to the externalist."[91]

It is not entirely clear just what Dancy's objection amounts to. There appear to be two ways to understand the objection. First, Dancy might be suggesting that an agent

---

[89] Dancy uses the concept of Lucifer put forth by John Milton in Paradise Lost.
[90] Dancy 1993, 6.
[91] Dancy 1993, 6.

cannot pursue evil for its own sake because that which is believed evil is something which is intrinsically repulsive to agents, while that which is believed good is intrinsically attractive to agents. Dancy seems to suggest this reading when he says that "for us to understand the pursuit of evil, we have to find some comprehensible relation between it and some good."[92] Second, Dancy might be arguing that by reflection on the very concept of a person who pursues evil, e.g., Lucifer, we realize that no one, not even Lucifer, actually pursues evil directly, but for extraneous reasons.

The first way of understanding Dancy's objection breaks down into two component parts. First, the claim that the good necessarily attracts and second that evil necessarily repels. I do not think that Dancy can claim that the good necessarily attracts given that he claims that such psychological failings like weakness of will, accidie, depression, and so on, can preclude an individual from pursuing what she believes to be good.[93] In response to the second claim Michael Stocker argues, this view is "clearly and simply false."[94] As Stocker explains: "there are clear and unproblematic [cases where we are attracted to do an act] because it is (believed) bad or in spite of its being (believed) bad, where the act or feature is not attractive because or only because it or some other relevant act or feature is (believed) good."[95] To illustrate Stocker's point, consider the following example: Imagine that Bill is suffering from AIDS and will soon die. He is angry and decides that if he has to die, then others should go with him. He begins to engage in sexual intercourse with people while neglecting to inform them that he has

---

[92] Dancy 1993, 6.
[93] For more on this type of response see Stocker 1979, 744.
[94] Stocker 1979, 740.
[95] Stocker 1979, 741.

AIDS. In sum: he has a desire to harm others. In this case, it seems that Bill wants to do something he clearly thinks is bad and despite the fact that it is bad. Now, it might be objected that Bill does not desire to do the harm intrinsically, but because it leads to something good, e.g., his own pleasure. However, as Stocker argues, "just as helping others can the direct and proper object of desires and appetites, so can harming others."[96]

How does Stocker justify this claim? Stocker puts forth the following:

> When we feel furious, hurt, envious, jealous, threatened, frustrated, abandoned, endangered, rejected, and so on, what we often seek is precisely the harm or destruction of someone, and not always the "offending party": "If I can't have her no one will." "So, you are leaving me after all I have done for you. Well than, take that." "The whole day has gone badly, I might as well complete it by ruining the little I did accomplish." "I let him have it with the horn; he was the millionth Sunday driver who cut in from of me." Watch out for him today, he just had an awful fight with his wife."[97]

Stocker's point seems to be that the mood I am in can influence my action, that is, "given such moods and circumstances [above], harming another can be the proper and direct object of attraction."[98] There does not seem to be any need to ascribe to the agent any other object which he is after, e.g., pleasure. In certain moods, I can desire what is harmful to others and myself. "In such moods, . . . we not only do not care, we are filled with 'uncare'."[99]

Now, if we accept that our moods can affect the connection between what we evaluate as good and what we are motivated to do, then we must accept that sometimes

---

[96] Stocker 1979, 748.
[97] Stocker 1979, 748.
[98] Stocker 1979, 748.
[99] Stocker 1979, 749.

when we are in certain moods we desire the bad. And, if we accept this, then we must reject the first reading of Dancy's objection to the possibility of an agent who pursues evil for its own sake.

Let us consider the second reading of Dancy's objection, *viz.*, that by reflection on the paradigm case of an evil-doer, e.g., Lucifer, we see that not even he pursues evil for its own sake, However, Dancy's response is problematic, since it can be used against him, *mutatis mutandis*, in the case of a person who putatively pursues good for its own sake (*sub specie boni*). Consider the paradigm case of a person who pursues good for its own sake: Jesus. The externalist might well wonder whether such an extreme degree of good-doing is actually possible for Jesus, or any person. For as we have seen it is sometimes the case that agents pursue evil. So perhaps the best way to understand Jesus's pursuit of good is to see Jesus as pursuing good, not for its own sake, but for extraneous reasons, e.g. because doing good will make it such that more people will be apt to believe and follow him, or because this is the only way to keep together his followers, or because it undermines the various activities of Lucifer.

If the externalist considers Jesus's actions in this way, then he can make sense of Jesus's pursuit of good: Jesus pursues good, not for its own sake, but instrumentally, i.e., because of what comes from it. This makes Jesus's pursuit of good explicable, but at the cost that it undermines Dancy's contention that a mere belief that something is good (or that we ought to do it) will motivate an agent to act. Moreover, what follows from this understanding of Jesus's pursuit of good is that it is possible to conceive that there is no person who pursues good for its own sake, or does what is right just because it is right.

People pursue good only for instrumental reasons. Accordingly, internalism is undermined, since it seems that no person is motivated by a moral judgement, e.g., 'It is right to x in circumstances c', just because it is right to do x. Instead, agents act on moral judgements because their actions satisfy some extraneous want or desire that an agent may have, e.g., the desire to do what is right. And, the externalist can readily accept such a conclusion. The internalist, however, cannot: he holds that a belief about what one ought to do motivates without the help of a desire, not because it leads to the satisfaction of some preference or desire.

Dancy might respond here by arguing that while no one ever acts on a moral judgement because it is right to do so, but only because it will, for example, mitigate the pain of a lover, or make a friend happy, this is only because "there may be no *general* reason for doing what one should."[100] Dancy argues that there are particular reasons, rather than one general reason, for acting morally, and, correspondingly, that there is more than one type of moral belief that can motivate an agent to act. Thus, an agent can be motivated to act morally by the belief that her action will cause a person pain or that her action will bring about the happiness of a child, and so on. Yet, my objection can be made to work here too. It may be held that no one will pursue something like the mitigation of another's pain for its own sake, but only for extraneous reasons, e.g., because an agent does not like pain or because he wants to make others happy, and so on. But the internalist cannot admit that people are motivated to mitigate pain only because they desire something which comes from doing so; the internalist holds that one is

---

[100] Dancy 1993, 10.

motivated to mitigate pain because in the circumstance in which the action takes place the stopping of pain is the morally correct thing to do, and this is why we are motivated to stop the pain. And, if the externalist may be allowed this move against the internalist, then it seems to demonstrate that agents are not motivated to pursue good for its own sake, but for other extraneous reasons. And, this contravenes internalist dogma.

It appears that Dancy is caught in a dilemma. If he argues that no one pursues evil for its own sake, then the same argument can be made for the case of someone who allegedly pursues good for its own sake. If Dancy wants to maintain that agents pursue good for its own sake, then he must concede that there could be some agents who pursue evil for its own sake. No matter which horn of the dilemma he takes, we get the same result: internalism is undermined. Internalism is undermined if Dancy takes the first horn, because it seems to be the case that no agent pursues good because he believes he ought to, but because he desires something which comes from it; absent that desire, the agent would fail to be motivated. Internalism says the opposite of this contention. Therefore, internalism is false. Moreover, internalism is undermined if Dancy takes the second horn of the dilemma, because on this horn it follows that there are agents who believe that what they are doing is evil, but pursue evil despite this knowledge. Internalism denies this thesis. Therefore, internalism is false.

I must be clear about my objection. I do not think that it undermines internalism because a desire happens to be present when an agent is motivated to act morally; an internalist can readily admit that one can be motivated to act in accord with morality and desire to do so. What my objection says is that it is possible that when agents act on

moral judgements the only reason they do so is *because* their actions satisfy some extraneous want or desire. Absent the desire, no agent would not be motivated to act morally. And the internalist cannot accept this result. Thus, Dancy's response to the possibility of the existence of the amoralist and the wicked person fails. Since his response fails, his argument for internalism is undermined. Externalism thus remains the stronger theory.

I think that there are two rejoinders to my argument here, but that both fail. We might interpret Dancy's objection to the wicked person as an objection against the possibility of there being any such thing as "pure evil." Rather, evil is simply the privation of good. There is no evidence that this is what Dancy does say, but he might suggest this as a defense given that if there is no such thing as pure evil, it would impossible for an agent to pursue it. Hence, it is impossible for there to be a wicked person as conceived, since a wicked person pursues pure evil. If there is no such thing as a wicked person, then the objection poses no problem for internalism. However, I think that this is a queer metaphysical doctrine. I think it is implausible, if not empirically false, to deny that there is such a thing as "pure evil." Hence, it seems implausible to deny that there are people who pursue evil for its own sake. Consider events like the Holocaust or mass warfare; *prima facie* these events are instances of positive evil in the world. Furthermore, I would deny that evil is defined in terms of good, such that evil is merely the absence of good. The claim that evil is merely the absence of good can be questioned by pointing out the symmetrical possibility that there is only evil in the world, and that good is nothing more than the absence of evil. To decide between these two alternatives,

one would have to rely on one's theological leaning, and so on. In other words, there may be no deep philosophical or metaphysical distinction between a world predicated on evil and a world predicated on good. Moreover, it seems to be the case that the argument against Dancy could be made to work even if I accepted the queer metaphysical doctrine that there is no such thing as pure evil. Instead of maintaining that there is a person who pursues evil for its own sake, I could maintain that there is a person who pursues the annihilation of good for its own sake. This postulation still poses a problem for internalism in the ways I put forth above.

The second rejoinder is this: my argument appears to be a non-sequitur.[101] I argued that perhaps we can only make sense of Jesus's putative pursuit of good for its own sake if we construe his pursuing of good not for its own sake, but for extraneous reasons. I claimed that it follows from this that an agent never pursues good for its own sake, i.e., a belief that something is good never motivates. But this is a non-sequitur. For it might be the case that a person does not always pursue good for its own sake, but it does not follow that one *never* pursues good for its own sake. Perhaps it is the case that one sometimes pursues good for its own sake, and sometimes one pursues good for extraneous reasons. The correct response here is to argue that even if it is the case that some people do pursue good for its own sake some of the time, then arguably people pursue evil for its own sake some of the time. This response is in keeping with my contention that the two cases are to be treated symmetrically. And, if it is true that some people sometimes pursue evil for its own sake, then internalism is false.

---

[101] Thanks to Margaret Cameron for bringing this objection to my attention.

3.3 Weakness of Will and The Person Suffering From Accidie

Dancy thinks that most cognitive theories of motivation fall prey to the problem of weakness of will and the problem of the person suffering from accidie.[102] Hence, if his cognitive account of motivation is to supersede these theories it needs to find an answer to these challenges. In this section, I will detail Dancy's response to these objections. I deal with his response to the problem of weakness of will first.

Internalism holds that if the total cognitive state of an agent is sufficient for motivation in one situation, then it must be sufficient for motivation in analogous situations. But the very same state can be present without generating action in the same circumstances when an agent is suffering from weakness of will. But, then, it must have been false that the state was sufficient for action in the first case. It must have been the case that there was something else, in addition to the cognitive state in the first case, which motivated the agent to act, namely, a desire. It seems that Dancy's theory is lacking something which explains action, e.g., a desire. Thus, it must be false. [103]

In reply, Dancy argues that this argument rests on a false premise, and thus is unsound. Dancy articulates the response this way: "[the argument] is unsound because it makes a basic assumption which the cognitivist [internalist] does not need to share, [namely,] that if a state is anywhere sufficient for action it must be everywhere sufficient."[104] Dancy's contention directly contravenes Hume's suggestion that there are states which are essentially motivating, i.e., desires, and states which are contingently

---

[102] Dancy 1993, 12.
[103] Dancy maintains that this argument is more a refutation "of cognitivism in the theory of motivation," (22) rather than a refutation of internalism.
[104] Dancy 1993, 22.

motivating, i.e., beliefs.[105] On Hume's account there are states which, when they are present, always motivate in the absent of contrary motivations, i.e., which cannot be present without motivating an agent to act, and states which contingently motivate, but never in their own right.

Dancy thus claims that there is nothing resembling beliefs and desires as conceived by Hume. Rather, there are "*intrinsically* motivating states, which can be present without motivating, but which when they do motivate do so in their own right. They can motivate in their own right, and so are not Humean beliefs; but they can be present without motivating, and so are not Humean desires."[106]

But what takes place when a state which is intrinsically motivating is present, but does not motivate? Dancy argues that "when this happens [when an intrinsically motivating state is present, but fails to motivate] there will be an explanation of it and we should expect to give this in terms of the presence of some feature in the second case which defused the ability of the original cognitive state to motivate."[107] Presumably, the explanation that Dancy has in mind will include reference to weakness of will.

Dancy's response to the problem of the person suffering from accidie is similar to the response to the problem of weakness of will. According to Dancy, "people who suffer from accidie are those who just don't care for a while about things which would normally seem to them to be perfectly good reason for action."[108] Accidie can be brought on by such things as depression. Accidie is a problem for the internalist because in such a case

---

[105] Beliefs are contingently motivating states because they motivate an agent to act only in the presence of a desire.
[106] Dancy 1993, 24.
[107] Dancy 1993, 24.
[108] Dancy 1993, 5.

an agent believes that she is under a moral obligation, but fails to be motivated by that belief. But internalism says that beliefs motivate; therefore, internalism must be false.

Dancy's response to this objection depends on the notion of an intrinsically motivating state. He replies as follows:

> the flexibility introduced by the notion of an intrinsically motivating state is surely just what we need to account for accidie. What we assert is that a state which is here sufficient for action may elsewhere not be. Where it is not sufficient, there will be an explanation for this. And we have introduced no restriction on the sort of explanation that we are prepared to countenance. Sometimes the reason will be carelessness or inattention; sometimes it will be despair; sometimes it will be an excess of alcohol; sometimes it will be a neurological disorder; and sometimes it will be clinical depression.[109]

Thus, if accidie can be included in what counts as an explanation as to why a state failed to motivate, Dancy's pure theory can quite easily accommodate the problem of accidie. Thus, for Dancy the problem of weakness of will and the problem of accidie are no problem for his internalism if he is allowed to posit the existence of an intrinsically motivating state.

I do not think that the Humean (the externalist) would disagree with Dancy's claim that weakness of will and accidie can prevent a state, which usually would motivate an agent to act, from motivating an agent to act. But neither do I think that this claim licenses a rejection of Hume's claim that there are essentially motivating states and contingently motivating states. The Humean can argue that, though a desire is an essentially motivating state, it can sometimes be present and fail to motivate an agent to act; but this does not entail that it is not an essentially motivating state. The failure of the

---

[109] Dancy 1993, 25.

state to motivate can be explained by some factor like weakness of will, accidie, and so on. The Humean can argue that it is true that people who are suffering from weakness of will and accidie are not moved to act by the reasons they take themselves to have, because cases where an agent is suffering from weakness of will, accidie, is drunk, and so on, *are* just those cases in which the agent fails to act in ways that they otherwise desire to act.

The Humean conception of rationality says that the rational thing for an agent to do is to maximize the satisfaction of her desires.[110] But when an agent is suffering from a psychological failing, she fails to act rationally; she fails to make it as likely as possible that she get what she most wants. By the agent's own lights she is irrational. We explain the agent's irrationality by appeal to some feature which is present in the case where the agent acts irrationally, but which was not present in the case where the agent acted rationally, e.g., weakness of will, accidie, drunkenness, and so on.

For the Humean, when an agent makes a moral judgement - and she is not suffering from weakness of will, accidie, not an amoralist, and so on - she will be motivated to act in accord with what the moral judgement enjoins, because she desires to do what the judgement prescribes. However, when an agent is weak-willed, suffering from accidie, or what not, her actions do not follow her desires. She fails to act on the reasons - beliefs and desires - she takes herself to have. Thus, by her own lights she is acting irrationally.

---

[110] See Gauthier 1975.

To illustrate my point, consider the following scenario. Imagine that a man's child is drowning in his backyard pool. From the window of his house the man sees the child, and he strongly desires to save the child from drowning. Further, he believes that if he leaves his house and jumps into the pool, he can save the child. But the man is suffering from agoraphobia and, hence, is unable to leave the confines of his house to save the child. Consequently, the child dies. What explains his inaction? Clearly in this situation the man suffers from a psychological failing, e.g., agoraphobia, and this explains why he did not act in a way that he had most reason to act.[111] Given this case, the Humean can agree with Dancy that certain psychological failings can preclude an agent from acting in ways that he has, by his own lights, most reason to act. The Humean, then, can agree that on this occasion an essentially motivating state failed to motivate. Moreover, if the father was not suffering from a psychological failing, he would have saved the child. When an agent suffers from psychological failings, his actions do not follow the perceived strength of the reasons he takes himself to have. The Humean can admit that essentially motivating states are motivating just if the agent whose states they are is not suffering from weakness of will, accidie, and so on.

---

[111] In this case, doesn't the man simply have an overwhelming desire to stay inside? I argue not. It is important to make a distinction here between having a strong desire to do x and being compelled to do x. In this case, the man is compelled to stay inside in virtue of the fact that he is afflicted with agoraphobia. No amount of convincing, including pointing to the fact that he strongly desires that his child not drown, will move him to act in accord with the reasons he has. His agoraphobia causes him to act irrationally. But perhaps he is compelled to have a desire to stay inside. If this is the case, then he has an irrational desire. His desire is irrational because it is based on a false belief which is caused by his agoraphobia, namely, that if he goes outside something bad will happen to him. His desire is irrational because it is caused by a belief which is itself irrational to hold given that desires that are caused by irrational beliefs are themselves irrational. Either way he acts irrationally: he fails to make it as likely as possible that he get what he most wants, namely, to save his child from drowning.

But perhaps my example misses the point.[112] Thus far I have assumed that what happens in cases where an agent is suffering from weakness of will, accidie and so forth, is that the agent's motivations are overridden by contrary motives. But Dancy's position might well be that what happens in these cases is that a person is simply not motivated at all. I think that there are two responses to this claim. First, if what I have said above is Dancy's actual intention, it does not appear clear what he could have in mind. For it is typically the case that when an agent suffers from weakness of will and accidie her motivations are typically overridden by contrary motives rather than simply truncated. Moreover, I think that Hume can accommodate Dancy's objection and, thus, demonstrate that it is false that there are no such things as essentially motivating states.

Consider again the case of the man who is suffering from agoraphobia. The way that I devised the case, the man's motivations were overridden by his psychological failing; but suppose instead that he was drunk. He sees that his child is drowning, but because he is drunk he fails to act on the reasons which he takes himself to have, i.e., what he believes is good is not what he is motivated to do. If Hume allows that there are such cases, as I think he can, then he can argue that the drunkenness cuts off the motivation, but that were the man sober he would have moved to save the child. Again, the state would be an essentially motivating case just if the man is sober. Thus, Hume can conclude that a state will be essentially motivating just if a person is not suffering from a psychological failing, not drunk, not sick, and so on. Thus, Dancy's contention that there is no such thing as an essentially motivating state is false.

---

[112] Thanks to Rich Campbell for pointing this out.

What Dancy has to demonstrate is that it is impossible for a person who is not suffering from weakness of will, accidie or other such psychological failings to believe that she has a reason to do something, and still not be moved to act, e.g., as in the case of the amoralist. But Dancy does not seem to be able to explain what happens in such situations, for the agent is not suffering from any psychological failing which could explain her failure to be motivated. I suppose Dancy might maintain that such a person could not exist, because if a person really did believe that she was under a moral obligation, then she would be motivated. But this begs the question against the Humean and the possibility that a person can believe she has a moral reason to do something, but not be motivated to act in accord with that reason. Thus, Dancy's response to the problem of weakness of will is flawed. It is here that externalism emerges as the stronger theory of motivation, because it can explain why an agent does not act in the way that morality enjoins; she fails to act in that way because she lacks a desire. Thus, externalism is explanatorily more potent than Dancy's internalism and emerges as a stronger theory.

In response to this objection, Dancy argues that a moral belief can fail to motivate an agent to act even when he is not suffering from weakness of will, accidie, and so on. Dancy argues that this follows from his claim that there are intrinsically motivating states. The notion of an intrinsically motivating state makes it possible to claim that "what are reasons [for action] *here* may not be the same reasons *there*, because of the presence of further reasons in the second case."[113] This putatively follows from the fact that "the ability of a consideration to motivate can be affected by background conditions

[113] Dancy 1993, 24.

which are not themselves motivators."[114] I suspect that what Dancy dubs 'background conditions' are just different non-moral properties; his point here is just that what is morally salient can change from situation to situation. In some situations a consideration will count in favour of an action. Further, in some situations a consideration will count against an action, while in others it will be irrelevant to how the agent ought to act. Of course, in the cases where the consideration fails to be a reason for action there will be an explanation of this fact. But this directly contravenes the (common) assumption that reasons are general in the sense that that if a consideration counts as a reason in one case, then it counts as a reason in every case. Let's call this assumption **The Generalist Assumption** (GA). In the next section, I defend GA against Dancy's objections. Further, I claim that since it is true that if a reason functions in one case to motivate action, then it functions to motivate in every case, the notion that there are intrinsically motivating states must be false, since it entails the possibility that the very same moral reason can function to motivate in some situations and not others, even if the situations do not differ in their non-moral properties.

3.4 Particularism Undermined

Dancy's argument in favour of the notion of an intrinsically motivating state entails his claim that GA is false. He takes his claim that GA is false to be a positive argument in favour of his normative theory, namely, particularism. I now examine this theory. I argue that particularism is false. If Dancy's particularism is false, then his claim that there are intrinsically motivating states is false. And, since this claim forms the foundation of his

---

[114] Dancy 1993, 24.

reply to the problem of weakness of will and accidie, his reply fails. Therefore, Dancy's account of motivation is undermined by the problem of weakness of will and accidie.

Dancy claims that we should accept that a belief-state can be sufficient to motivate action on one occasion but not in another, without adding a further state to explain action, e.g., a desire. For instance, imagine that I believe that action x will increase pleasure. Dancy suggests that this belief might motivate on one occasion, not motivate on another. According to Dancy, if we accept the notion of an intrinsically motivating state, it follows that moral reasons function in the same way: the same consideration may count as a reason for doing an action in one case, yet not count in an another. Thus, that something increases pleasure can function as a reason for doing an action in one case, but perhaps not in another situation. This is Dancy's normative theory, which he calls particularism. Dancy claims that particularism is the view that are no general moral rules or "substantial moral principles."[115] He articulates the core thesis of particularism this way: "a property F of one action may be a reason for me to do that action, even though the F-ness of another action may be morally indifferent or even count as a reason against doing it."[116] He links up particularism and the notion of an intrinsically motivating state this way: "What we see here is that changes in the attendant circumstances can alter rather than overwhelm the moral tendency of a particular property. This is just like the ability of other mental states . . . to alter rather than to

[115] Dancy 1993, 66.
[116] Dancy 1993, 55-56.

overwhelm the motivation of a motivating state. It is the crucial respect in which the logical behaviour of reasons is like that of morally relevant properties."[117]

The most appropriate way to demonstrate the core thesis of particularism is by example. Imagine that the proponent of GA asserts that pleasure is always a reason for doing an action. But Dancy argues that in many cases it can be a reason against doing an action. For example, that my friends and I receive pleasure from being taken to a movie can constitute a reason for taking my friends out for a movie, but that I receive pleasure from torturing young babies can count as a reason against my action. Translated into moral psychology, Dancy's claim amounts to the suggestion that the belief that my action increases pleasure can in one case, e.g., that of taking my friends to a movie, constitute a reason for me to act, but the very same belief can constitute a reason against my action in another case situation, e.g., that of torturing babies. Thus, it must be false that a moral reason functions in the same way in every situation, and false that a belief representing that moral reason functions in every case to motivate an agent to act.

How might the generalist reply to this objection? A possible response could be to argue that we can always alter general reasons for action in the following way: 'Pleasure counts as a reason for action except in situation x, y, z.' If this was the rule, then it would still be a general reason for doing an action. Thus, in all cases, except x, y, and z, pleasure counts as a reason for action. Further, in every case, except x, y, and z, the belief representing the moral rule motivates an agent to act, except in circumstances, x, y, and z. The generalist can always simply revise and improve her principles to meet with the

---

[117] Dancy 1993, 56.

objections Dancy puts forward. If the generalist makes this move, then she can claim that a reason does function as motivation to act in all circumstances except , x, y, z, and hence that a belief state embodying those reasons functions to motivate action in all situations except x, y, and z. If this is the case, then particularism about moral reasons is false; and, if it is false, then it is false that there are intrinsically motivating states, since that there are intrinsically motivating states entails that moral reasons are not general.

There are two objections to this proposal. First, perhaps this reply commits the generalist to an implausible account of general principles in the sense that there would be so many extenuating circumstances, e.g., so many circumstances like x, y, and z, generating so many principles that it would be impossible for one agent to remember them and act on them. That is, the principles in question may become so specified that they would be too detailed to be useful to the agent.[118] But this is hardly an objection to GA. Particularism falls prey to the same problem. Dancy suggests that the way to identify moral reasons in certain cases is to identify what is salient in a certain case.[119] But an agent may find it equally hard to discern what is morally salient in each case, since there is no principle to which the agent can appeal which can discern for him what is the right thing to do in each case. Further, it might well be equally difficult for the agent to pick out which property would constitute a moral reason and a reason for action. So both GA and particularism fall prey to the same problem

The second objection to the above proposal is this. Dancy might argue that by adding reference to particular cases a generalization is particularized. This would concede

---

[118] McNaughton 1988, 197.
[119] Dancy 111-116.

to the particularist that there is nothing that all right or wrong acts have in common. *Prima facie* this looks like a reply, but on closer inspection it really isn't. This follows from the fact that particularism denies that there are moral rules or substantive moral principles. But in the cases where the moral rule is rectified to apply only to certain cases, it still remains the case that in all situations, except x, y, and z, there is something that all these acts have in common, e.g., they either increase or decrease pleasure. There is nothing prohibiting the generalist from constructing another rule which applies to cases x, y, z, and, hence, nothing stopping the generalist from saying that there is something - the same thing - wrong or right about all cases x, y, z. Further, a general rule need not be a rule which covers all acts. A general rule may be very specific and cover only particular cases, but it is still a general rule that is to be followed in particular cases. This is something the particularist denies.

However, a better reply to Dancy on behalf of GA can be generated from closer examination of the above cases. In the first example, that my friends received pleasure from being taken to a movie counted as a reason for taking my friends to a movie, but that I received pleasure from torturing babies counted as a reason against my torturing babies. What Dancy seems to be objecting to here is not so much that pleasure counts as a reason for action in every case, but that pleasure can count as a reason for action in cases where one is deriving pleasure from something which is obviously (intuitively speaking) immoral. The exponent of GA might reply by saying that pleasure counts as a reason for action in all those cases where pleasure is not derived from an immoral act. What counts as a reason for action in the generalist view is not pleasure *simpliciter*, but pleasure from

what is not immoral. Further, our rejection of the moral wrongness of deriving pleasure from torturing babies seems to depend less on the circumstances in which the action takes place, which is something on which Dancy places great emphasis, and more on the fact that torturing babies is just plain wrong, and there is a general moral rule against it. Turning to moral psychology, the promulgator of GA could say that a belief that an action gives pleasure motivates one to act only in those cases where the pleasure is not derived from an immoral act. Thus, this belief will be a reason for action in all cases where the action is not immoral. This would buttress GA from Dancy's objection and enable those who espouse GA to claim that if a moral reason functions as a reason for action in one case, then it serves as a reason for action in every case (except where the action is immoral). So pleasure counts as a reason for action in all cases, but pleasure from what is immoral counts as a reason against action. The same can be said for beliefs and motivation.

There is another reply to Dancy's objection to GA. Imagine that it is a general rule that 'It is wrong to lie in circumstances x'. What follows from this rule is that whenever one finds one's self in circumstances x, one ought not to lie. But this is a general rule which is impervious to Dancy's claim that the same reason for action does not always function as a reason for action, given differing circumstances. Dancy cannot claim that the altered circumstances alter the moral reason, because the circumstances are identical in their non-moral properties, and so on. And, if we take it that a moral property is fixed or fashioned by its non-moral properties, then if a situation is identical in all its non-moral properties, it will be identical in all its moral properties. Alternatively, if the non-

moral properties (i.e., the facts which comprise the circumstance in which I act) of a situation differ, then the moral properties differ. Again, the exponent of GA can argue that a belief that lying is wrong can function as a reason for action in all x-like cases. Thus, Dancy's objection to GA can be met. And if it can be met, Dancy's claim that a moral reason does not always function in the same way must be false; hence, it must be false that a mental state can be present in one case and motivate, yet be present in another case and not motivate, since this entails particularism about moral reasons, and that doctrine is false. If this is this so, then Dancy's attempt to justify his argument against weakness of will and accidie fails.

3.5 The Argument From Directions of Fit

In the last chapter (2.2), I reported McNaughton as arguing that a single state could have the direction of fit of both a belief and a desire. He made this suggestion in order to avoid falling prey to the argument from directions of fit. According to this argument, the best way to distinguish between beliefs and desires is to appeal to the notion of directions of fit. On this view, beliefs and desires have different and opposing directions of fit. Beliefs function to represent the world, and their being true is their fitting the world. The direction of fit of a belief is the mind-to-world direction, i.e., a belief state is one which must fit the world. Desires do not function to represent the world as it is; they function to move agents to act in accord with the content of their desires. A desire has the world-to-mind direction of fit, i.e., it is a state with which the world must fit. This is why desires move agents to act. Dancy expresses the distinction this way: "a belief is a state which

aims to be caused by the truth of its content, while desire is a state which aims to cause its own content to become true."[120]

But this way of distinguishing beliefs and desires poses a problem for Dancy's cognitivist internalism in the following way. Without the presence of two distinct states, each with a different direction of fit, action is impossible. In order for an agent to be motivated to act she requires both a belief and a desire. But Dancy's claim that two belief representations, one representing how things are, the other representing that a certain action would change those things in a certain way, is enough to motivate an agent to act, does not ascribe to an agent a state with the direction of fit of a desire. Thus, it does not explain why an agent acted.

According to Dancy, what forms the basis of the claim that a motivating state must be comprised of two different states with two different directions of fit is a "Cartesian metaphysics."[121] The main idea behind this view is that "the mind and the world are radically independent and therefore [this view] conceives of the world as intrinsically inert."[122] The world is inert in such a way that it "contains nothing important or relevant to action and choice."[123] Dancy suggests that if we accept this view, we have to accept that "for there to be action there must be a desire present, for rational action is inconceivable without a state which motivates one to make the world be the way [that desire] conceives of it being."[124]

---

[120] Dancy 1993, 28.
[121] Dancy 1993, 31.
[122] Dancy 1993, 31.
[123] Dancy 1993, 31.
[124] Dancy 1993, 31.

Dancy rejects this metaphysical view and with it the notion of direction of fit. His pure theory is committed to the claim that an agent's beliefs can, in and of themselves, motivate her to act. But what Dancy is suggesting with this latter claim is that "what motivates is the matter of fact believed, not the believing it."[125] For Dancy, facts, not beliefs, motivate agents to act, and this directly contradicts the claims made by Cartesians. On Dancy's view, to say that beliefs are what intrinsically motivate is just to say that "that there are facts which intrinsically make a difference to how we should act."[126] If we assert this, then we have abandoned the Cartesian metaphysical view. And, since the Cartesian metaphysical view cannot allow for the claim that there are "facts which intrinsically make a different to how we should act,"[127] both the metaphysical view and the notions of directions of fit which reject it are flawed.

This is an ingenious rebuttal to the argument from directions of fit. However, I think that it fails. Dancy began his discussion of motivation by arguing over what is the best account of the considerations which motivate action and so *explain* action. However, this last argument is no longer tackling this subject; it has moved on to discuss what *justifies* action. Thus, there is a distinction between what motivates and explains action and what justifies action. It is his failure to respect this vital distinction which forms the basis of my reply to this objection.

Brink argues that the phrase 'reasons for action' can be read in two different ways. First, we can make reference to an agent's reasons for action in an attempt to articulate

---

[125] Dancy 1993, 32.
[126] Dancy 1993, 32.
[127] Dancy 1993, 32.

what motivates an agent and so explains her action. According to Brink, "here we use 'reasons for action' to refer to the considerations that *motivate* an agent and so *explain* her actions."[128] Second, we can read 'reason for action' as "*good* or *justifying* reasons."[129] An agent can have explanatory reasons for action without having justifying reasons.

Brink gives the following example to help illustrate the distinction:

> if I am a light-bulb eater, my belief that light bulbs are nutritious and my desire to be healthy constitute my reason, in the first explanatory sense, for eating light bulbs. Although presumably I do not have a reason, in this second justificatory sense, to eat light bulbs.[130]

This example helps illustrate where Dancy has gone wrong in his objection. It is important to separate the agent's belief from the content of the agent's belief. But Dancy fails to do just this. In the above case, to explain an agent's action we appeal to her beliefs and desires, such that if we asked the agent 'Why did you eat the light bulb?' she would answer 'Because I believed that it was good for my health, and I desire to be healthy.' What motivates the agent and so explains her action is her belief and her desire, not a fact. So in explaining an agent's action we must make reference to her beliefs and desires. As Michael Smith points out: "The distinctive feature of a motivating reason to Φ is that, in virtue of having such a reason, an agent is in a state that is explanatory of her Φ-ing . . . Given that an agent who has a motivating reason to Φ is in a state that is in a way potentially explanatory of her Φ-ing, it is thus natural to suppose that her motivating reason is itself *psychologically real*."[131] Furthermore, Smith continues, "it would seem to

---

[128] Brink 1989, 39.
[129] Brink 1989, 39.
[130] Brink 1989, 39
[131] Smith 1994, 96.

be part of our concept of what it is for an agent's reasons to have the potential to explain her behaviour that her having those reasons is fact a about her; that the goals that such reasons embody are her goals."[132] Motivating reasons are thus psychological states, "states that play a certain explanatory role in producing action, [rather than facts]."[133]

By contrast, in the case in which we are trying to surmise what could possibly be justification for an agent's action, we are not going to look at her psychological states (i.e., facts about that agent); rather we are going to look at what facts may justify her action. Smith calls justifying reasons normative reasons.[134] On Smith's account, "to say that someone has a normative reason to Φ is to say that there is some normative requirement that she Φs, and is thus to say her Φ-ing is justified from the perspective of the normative system that generates the requirement."[135] For Smith, normative reasons are best thought of as facts or propositions about what we ought to do.

How does this distinction undermine Dancy's objection to the argument from directions of fit? If the Humean plays up this distinction, then it seems that he can reject the Cartesian metaphysical outlook according to which the world is inert, and hold that there are facts which justify our actions, but still retain the claim that an agent's actions are explained by appeal to her psychological states, namely, beliefs and desires.[136] If the Humean does this, then he is able to still appeal to the notion of directions of fit to explain why an agent acted.[137] And, if this is so, then Dancy will need to find an answer

---

[132] Smith 1994, 96.
[133] Smith 1994, 96.
[134] Smith 1994, 95-96 &131-132.
[135] Smith 1994, 95.
[136] Smith does just this, but more on that in the next chapter.
[137] How can the Humean have it that mere facts can justify action? For my purposes, I do not think that this question really needs answering. I need merely assert that it is a possibility. However, Smith does give an

to the claim that what motivates and so explains action is that an agent has a desire and a belief, that is, two states having different and opposing directions of fit. He has failed to do this, so we should reject his account of motivation for failing to ascribe to the agent states with the correct direction of fit to explain the agent's action.

But perhaps Dancy's point is that a fact can both justify *and* explain action. Agents can and do deliberate on what they believe is right and in virtue of this deliberation be moved to act. In order for this reply to help Dancy out it would have to be the case that agents always act in accord with what they believe is right, but this would beg the question against the arguments from direction of fit. The argument from directions of fit requires that for an agent to be moved to act an agent must possess both a belief and a desire, i.e., two states with different directions of fit. All Dancy's reply says is that an agent is moved by a state with the direction of fit of a belief and, hence, this is really no answer. Further, agents sometimes act contrary to what they believe is right, e.g., when they are suffering from weakness of will, accidie, and so on. It is not clear how Dancy could explain what is motivating agents to act in these situations. But perhaps there is another way for a fact to both explain and justify. For example, the fact that Φ-ing is right can explain why I come to believe that Φ-ing is right. This belief can in conjunction with my desire to do what is right explain why I Φ. Thus, a fact, in cases where I have access to the fact and have a desire to do the right thing explains and justifies my Φ-ing. This is a legitimate move, but it plays right into the Humean's hands,

---

answer to this question. He argues that there is a fact of the matter about what it is rational to desire. It is rational to desire what all fully rational, epistemically privileged agents would desire in a given situation. If an agent desires otherwise, then his action is not justified, though it can still be explained by reference to beliefs and desires.

for it suggests that a desire must be present for an agent to act, that is, it says that a state with the direction of fit of a desire must be present to motivate and so explain and agent's action. But this is just the Humean's motivation view. Moreover, this move would have the consequence of undermining Dancy's entire theory because he denies both that a desire needs to be present for an agent to have reason for action and that a desire plays a causal role in action.

It appears that what Dancy really objects to here is not the Humean account of action explanation, but the Humean account of normative reasons, or reasons which justify our actions. According to the Humean conception of rationality, it is rational to make it as likely as possible that one will get what one most desires. That is, it is rational for you to maximize the satisfaction of your currently held desires. This is the 'maximizing' conception of rationality.[138] Hume thought that what motivates and so explains action is that an agent has both a belief and a desire. Moreover, Hume thought that the having of a desire is what justifies an action. Dancy can reject Hume's account of normative reasons, but this rejection does not entail the rejection of Hume's account of motivating or explanatory reasons. Thus, even if we grant Dancy's objection to the Humean account of normative reasons, Hume can still hang on to his account of explanatory or motivating reasons. And, if this is the case, then Dancy falls prey to the argument from directions of fit.

Dancy might try to reply by arguing that what motivates action is two belief representations, and that what explains why an agent acted was that he had these two

---

[138] See Gauthier, 1975.

beliefs. But this really is no reply at all, since the Humean asks why this should motivate an agent to act and so explain why an agent acted. And we find that no answer on Dancy's behalf is forthcoming. Dancy does say, against the Humean, that facts justify our actions. However, this is no objection to the Humean because as I have argued above, the exponent of the Humean account of motivation can maintain with Dancy that what justifies an action is that there is a matter of fact about what you ought to do. What the proponent of the Humean account of motivation rejects is that a fact can explain why an agent acted. If the exponent of the Humean conception of motivation accepts that facts can justify, but not motivate action, then the this type of Humean can accept Dancy's criticism and still maintain that what motivates and explains action is a combination of beliefs and desires. Thus on the Humean's motivational view, one can be aware of a fact that justifies an action, but not be motivated because one lacks a desire which is necessary for action. The burden of proof thus rests squarely with Dancy; he must show us how it is that two belief representations can motivate and so explain action.

Obviously Dancy cannot claim that facts by themselves motivate, since an agent can be ignorant of a fact. A fact can motivate only if an agent is aware of it or has a belief which represents that fact. And, if this is true, then beliefs will have to enter into the picture when explaining action, though facts can enter into the picture in justifying action. Further, if beliefs somehow motivate action and, hence, explain action, then Dancy has to show how this is the case, over and above rejecting the Cartesian metaphysical picture. If, as I have argued, the Humean can reject Cartesian metaphysics, then the he can still quite plausibly ask Dancy what motivates and explains action. Dancy cannot simply assert that

beliefs have this job, because that would be to fail to answer the very question which required Dancy to address the argument from directions of fit in the first place.

There is one more point that needs to be made clear. When the Humean agrees with Dancy that facts can play a role in justifying action, he is not saying that what motivates and explains action is a fact of the matter. The proponent of the Humean conception of motivation is merely saying that facts justify actions, but have no ability whatsoever to motivate nor explain action. What motivates and explains action is a combination of beliefs and desires. Dancy has failed to show us that this is false. Therefore, we should reject his cognitive internalism on the grounds that it fails to provide a plausible answer to the argument from directions of fit.

3.6 Further Problems for the Pure Theory

In this section I want to demonstrate that Dancy's pure theory, in addition to its inability to meet Humean-type objections, is problematic on its own terms. There is something not quite right with the claim that what motivates an agent is the "gap" between the two representations. For what is it about that gap which motivates an agent to "close" it? Surely Dancy cannot say that it is a desire, because that would collapse his theory into Humeanism. But then what is it that moves me from, for example, believing that my girlfriend is in pain and that if I give her the morphine, her pain will cease, to relieving that pain? Perhaps it is that I prefer that she not be in pain, or that I want to show her my love, or that I want to be a nice person. Yet Dancy cannot appeal to such claims because they are all desires, and he thinks that it is not desires but the two representations or beliefs that motivate me to act. Further, he does not want to suggest that the second

representation is a desire, since that would completely undermine his case against Hume; that is just what Hume holds. But then the reason why I am moved to close the gap remains a mystery! Dancy cannot even suggest that I do it because I want to do what is right, for that again would be an appeal to a desire. It must, then just be the mere fact that I believe that my girlfriend is in pain, and that I believe that I can, if I act, relieve that pain. But why should these beliefs motivate me? I have suggested several answers to these questions, all of which Dancy must reject, since they force him to embrace some kind of Humean account of desire. Thus it remains a complete mystery why I am motivated to "close" the gap between the two representations.

Now, Dancy does suggest that without a desire action is impossible, but for logical, rather than causal reasons.[139] So perhaps Dancy can appeal to his own concept of desire to explain why an agent was motivated to close the gap between two representations. But this Dancy cannot do given that on his theory it turns out that a desire is caused by, not the cause of, motivation. It would seem odd to say that one's desire motivated one's action and was caused by one's being motivated to act. Thus, this provides us with no answer to what in fact motivates agents to close the gap between two representations. Moreover, to concede that a desire motivated action would be to veer dangerously close to collapsing into Humeanism about motivation.

Dancy might attempt to reply here by arguing that being motivated is to be in a state of desire; but what explains being in a state of desire and being motivated is the cognitive gap between the two representations. I am not inclined to accept this reply.

---

[139] Dancy 1993, 20.

Recall that Dancy denies outright that a desire is causally required for action and that a desire is needed to explain why an agent acted. To claim that being motivated just is to be in a state of desire, then, appears to render the desire-state entirely superfluous to his account of action-explanation; moreover, the state of desire which, he alleges, is concomitant with being motivated, appears to play no causal role in an agent's being motivated to act. Therefore, either the state of desire is altogether superfluous for explanation or it is causally irrelevant.

Dancy is also faced with difficulties in cases where there is more than one second representation which related to a first representation. Consider the case above where I believe that my girlfriend is in pain and I believe that if I were to give her the morphine her pain would be mitigated. Imagine that I also have a second belief representation which looks like this: I believe that if I rub my girlfriend's lower back, then her pain will be mitigated. In this case there are two second representations. What in Dancy's theory can explain how I pick between the two representations in action-selection? It certainly cannot be that I desire one over the other.

The problem is exacerbated by the fact that, according to Dancy, the two representations are really no different from each other. Dancy argues that we should see the content of the second belief representation "as that of the subjunctive conditional 'If I were to act in such and such a way, this would be the result.'"[140] The two representations are in some way analogous because they have the same consequent, *viz.*, my girlfriend's pain is mitigated. For Dancy, when an agent acts, she acts to make the consequent of the

[140] Dancy 1993, 28.

subjunctive conditional true, but not the subjunctive conditional itself.[141] What seems to follow from this, though, is that *how* my girlfriend's pain is mitigated does not matter, since I am not acting to make how I am mitigating the pain true. All that I am acting to make true is the consequent of the subjunctive conditional which represents the content of the second representation, i.e., content of the belief. But there are difficulties here. Suppose I believed that were I to kill my girlfriend, then her pain would be relieved. Dancy argues that I am acting to make the consequent of this subjunctive conditional true. But there are two actions which I could take to make the consequent true, and it does not seem to be the case that Dancy has given a clear directive to the agent in order for her to decide which action she should take to make the consequent true. If this is the case, then there is no way to decide which second representation (i.e., which subjunctive conditional) is the one upon which I should act. Dancy fails to provide us with such a decision mechanism, and so his notion that the gap between representations motivates me to act must be flawed.

Notice that Dancy cannot claim that the latter second representation that I suggested is ruled out because it is obvious that I would not desire the death of my girlfriend. He cannot concede this because this would be to concede that a desire played a part in deciding which of the second representations was the one upon which I acted. And if he did this, it would no longer be the case that it was a belief alone which motivated me to close the gap. Dancy might attempt to suggest that which representations motivate me to act are the representations which represent the morally best state of affairs. Returning

---

[141] Dancy 1993, 28-29.

to the case above, we can assume the situation in which I give my girlfriend the morphine is the morally best state of affairs, rather than the situation in which I kill her to relieve her pain. This is an appropriate response, but I do not think Dancy is out of trouble here. Imagine that there are two morally best second representations. In the situation above suppose I have the following two beliefs: (1) If I give my girlfriend the morphine, her pain will be relieved. (2) If I give her some opium, here pain will be relieved. For the sake of argument suppose that these two representations represent equally "morally best" states of affairs - there is a tie for "morally best". How can Dancy explain which representation motivates me? He cannot say that the morally best one does, because there are two morally best ones. But this is his last line of defense against the objection that it remains a mystery what moves me to close the gap between my two representations.

Of course, Hume's dictum is no better in cases where an agent has two equally strong desires for some state of affairs. But Hume's account of motivation remains the stronger theory because it can explain why the gap between two representations is salient to an agent and, hence, why an agent would be moved to close it. Dancy's account of motivation according to which an agent is motivated to act by the gap between two representations, has trouble explaining why an agent regards the gap between two representations as salient to her, and why she is motivated to bring it about that the gap is closed.

3.7 Conclusion

If what I have said about Dancy's pure theory is correct, then we should reject it *tout court*. Like McNaughton, Dancy has failed to justify his theory beyond its initial intuitive

appeal. But if this is the case, then we really have no justification for buying into it. Both McNaughton and Dancy attempted to maintain that the best solution to the moral problem I sketched in the first chapter is to reject the Humean account of motivation and argue that a moral judgement expresses a belief about the way the world is, morally speaking, and that a moral judgement can motivate an agent to act. I have argued against this view that either it is false or that the Humean account of motivation can be defended against its objections.

In the next chapter, I take up a rather different view. This is the view of Michael Smith. Smith argues that the best way to solve the moral problem is to find a way in which one can hold simultaneously that moral judgements express beliefs, that moral judgements motivate agents to act, and that a complete motivating state is comprised of both a belief and a desire. I now turn to Smith's theory.

CHAPTER FOUR

Michael Smith: Cognitivism, Internalism and Humeanism at Once

In chapter one I sketched what Michael Smith calls "The Moral Problem". It can be expressed succinctly in the form of the following inconsistent triad:

> 1. Moral judgements of the form 'It is right that I Φ' express [only] a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.
>
> 2. If someone judges that it is right that she Φs then, *ceteris paribus*, she is motivated to Φ.
>
> 3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences.[144]

The moral problem arises because it seems *prima facie* impossible to hold all three statements consistently. In the previous two chapters, I canvassed the views of David McNaughton and Jonathan Dancy with respect to the solubility of the moral problem. Both argue that an intuitively plausible solution requires rejecting 3 while embracing 1 and 2. I have maintained that either they cannot effectively rebut 3, or that their theories themselves are flawed in their own right. I tentatively concluded that there is something fundamentally wrong with their ways of solving the problem. In this chapter, I discuss Michael Smith's putative solution. Smith, unlike McNaughton and Dancy, argues that the best solution to the moral problem requires demonstrating a way in which all three propositions can be held consistently.

Smith thus holds that morality is objective and practical, and that motivation requires a desire. His theory is at once cognitivist, internalist and Humean in its

---

[144] Smith 1994, 12.

conception of motivation. He attempts to establish this view by arguing that Hume is correct to claim that belief and desire are distinct existences which motivate and so *explain* action. But Hume is wrong to think that what it is rational for us to do is to make it as likely as possible that we get what we most desire; that is, Hume's account of rationality is false. Smith thus constructs an anti-Humean account of rationality. Smith's solution to the moral problem turns on his reinterpretation of 2. A moral belief is a belief about what we have reason to do, and the connection between beliefs and motivation is established by Smith's own theory of rationality. In this chapter, I argue that Smith fails to give convincing reasons for why we ought to accept internalism, and that there is something very wrong with his anti-Humean account of rationality.

4.1 The Practicality Requirement

Broadly speaking, internalism is the view that there is an internal or conceptual connection between an agent's believing or recognizing that she is under a moral obligation and being motivated to act in the way enjoined by that obligation. In other words, if an agent believes that she is under a moral obligation, she will be motivated to act morally. Several different claims fall under the rubric of internalism. Smith defends two. First, there is the claim that "If an agent judges that it is right for her to Φ in c, then she is either motivated to do Φ in c, or she is practically irrational."[143] Smith calls this the practicality requirement.[144] The second form of internalism Smith calls rationalism: "If it

[143] Smith 1994, 61.

[144] On the practicality requirement the connection between the agent' s belief and her motivation is defeasible because the connection can be severed when the agent is suffering from weakness of will, and the like.

is right for agents to Φ in circumstances c, then there is a reason for agents to Φ in c."[145] On Smith's view, rationalism entails the practicality requirement in virtue of the fact that if an agent has a reason to do Φ in c, then an agent is motivated, if rational, to do Φ in c (where one's being rational is understood as one's acting in accord with the reasons one has). Smith defends both of these claims against David Brink[146] and Philippa Foot[147]. Brink takes issue with the practicality requirement, while Foot dissents to rationalism.[148]

Exponents of the practicality requirement argue that if an agent believes that it is right for her to Φ, then she is motivated to Φ. David Brink argues that the practicality requirement is false, for there are people - for example, amoralists - who judge it right to Φ, but fail to be motivated by that consideration. [149] The possibility of the amoralist is inconsistent with the practicality requirement. Therefore, the practicality requirement is false. In response, Smith argues that although the amoralist attempts to make moral judgements, she fails to do so. But Smith does not think that this decides the issue, since it appears to be the case that there is no non-question begging way to decide whose concept of the amoralist is right.[150] But Smith claims not to be troubled by Brink's objection since he has a separate argument which establishes the truth of the practicality requirement and, correspondingly, the falsity of the claim that there are amoralists.

According to Smith, it is a "striking fact" that there is a reliable connection between a good and strong-willed agent's changing her moral views and her changing her

---

[145] Smith 1994, 62.
[146] Brink 1989.
[147] Foot 1972.
[148] It is important to point out that the claims that Smith endeavors to defend are conceptual claims, rather than substantive claims. See Smith 1994, 63-66.
[149] Brink 1989, 45-50.
[150] Smith 1994, 68-71.

motivation. The reliable connection manifests itself in the fact that "a *change in motivation* follows reliably in the wake of a *change in moral judgement*, at least in a good and strong-willed person."[151] To illustrate this, consider the following scenario. Imagine that Bill and Ron are arguing over the moral permissibility of abortion. Bill believes that abortion is fundamentally wrong and is motivated to bring it about that there are no more abortions. Ron believes that it is morally permissible to allow women to procure abortions and is motivated to bring it about that women be allowed to procure abortions. After a long conversation, Ron convinces Bill that it is morally permissible to allow women to procure abortions. According to Smith, if Bill is "a good and strong-willed person a new motivation will follow in the wake of [his] new judgement."[152] Since Bill no longer believes that abortion is wrong and now believes that abortion is morally permissible, he will now be motivated to bring it about that women be allowed to procure abortions. Smith suggests that since there is a reliable connection between an agent's moral judgement and his motivation, both promulgators of the practicality requirement and externalists alike will have to explain how this is.

There are two ways to explain this fact about the reliable connection. Those who accept the practicality requirement can explain the reliable connection *internally* by appeal to the content of the moral judgement alone. On this view, since an agent who believes that it is right to Φ will be motivated to Φ, it comes as no surprise that a change in moral belief is followed by a change in motivation in a good and strong-willed person. Thus, as Smith puts it, "if an agent judges it right to Φ in C, and she has not derived this

---

[151] Smith 1994, 71.
[152] Smith 1994, 72. Smith asks us to assume that Bill is a good and strong-willed individual.

judgement from some more fundamental judgement about what it is right to do in C, then, absent weakness of will and the like, the defender of the practicality requirement can insist that she will be motivated non-derivatively to Φ in C. This is because, [on Smith's view], a non-derivative desire to Φ in C is what her judgement that it is right to Φ in C causes in her."[153]

Externalists, of course, cannot argue that a moral belief or the content of a moral obligation is what motivates an agent to act, since they hold that it is something external to moral judgements which motivate agents to act morally. Smith suggests that the only way for the externalist to explain the reliable connection between moral judgement and motivation is to appeal to the "motivational disposition [agents have] in virtue of which [they] count as a good person."[154] That is, externalist will have to explain the reliable connection by appeal to the content of an agent's moral motivation. According to Smith, the content of an agent's moral motivation on the externalist's view will consist in a "motivation to do the right thing."[155] The reliable connection is explained by the fact that agents have a non-derivative desire to do what is right, and it is in virtue of having this disposition that when an agent changes her moral views her motivation changes.

Smith argues that there is something wrong with the externalist's proposal. Consider the case of Bill and Ron described above. Smith argues that Ron's motivations are "derivative because they are derived from [his] current judgement about what the right thing to do is together with [his] basic moral motive: a non-derivative concern to do

---

[153] Smith 1994, 73.
[154] Smith 1994, 73.
[155] Smith 1994, 74.

what is right."[156] Ron cares only derivatively for the things he thinks are right in virtue of the fact that he has only one non-derivative concern: to do what is right. All other desires flow from this concern. However, Smith argues that there is something very wrong with the externalist's proposal: it attributes the wrong content to the motivational dispositions possessed by a good and strong-willed agent. The externalist is committed to the claim that an agent cares only derivatively for the things he thinks are right. But as Smith explains:

> Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, and the like, and not just one thing: doing what they believe to be the right thing . . . Indeed, common sense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue."[157]

The only way that an externalist can explain the reliable connection between an agent's moral belief and his motivation is to put forth a false account regarding a good person's motivations, thereby elevating a moral fetish into the one and only moral virtue. Thus, externalism should be rejected. The explanation put forth by Smith faces no such difficulties, so it is to be preferred to externalism.

One might find Smith's objection to the externalist's account of the explanation of the reliable connection between moral judgement and motivation less than pleasing. It merely begs the question against the externalist. The externalist claims that the motivational disposition of an agent is that he desires non-derivatively to do what is right, while Smith thinks that the motivational disposition which we ascribe to an agent says

---

[156] Smith 1994, 74.
[157] Smith 1994, 75.

that an agent is motivated to do the right thing because he cares about these things for their own sakes; doing these things does not satisfy some other desire that he has, e.g., the desire to do what is right. The externalist might say that his account of the motivational dispositions of an agent is intuitively plausible, while the practicality requirement exponent denies this. How may we decide between these views?

The fact that we typically think that a good-person is one who cares non-derivatively for what he believes is right does not produce an argument in favour of the internalist position. It looks like we are merely battling over intuitions regarding the plausibility of the content of the motivational dispositions we ascribe to agents. Further, it is not clear that Smith is on the right track here. For we often describe an agent as doing things not for their own sakes but because of some other reason. For example, imagine a religious man named John. John loves his wife and children and attends church weekly. One day at lunch John is having a conversation with a co-worker. He asks John "Why do you care for your children and wife?" Suppose John answers him by saying: "Because that is what God commands." John's co-worker might consider his response strange, but I doubt he would think that John was not a good person. On Smith's view, however, John would fail to be a good person because he fails to care non-derivatively for the well-being of his family. For example, John does what is right because he has a non-derivative desire to do what God says he ought to do. Further, such people as Jesus, Mother Theresa, and so on, would all fail to be good people, because they attempt to mitigate world hunger, and so on, not because they care non-derivatively for these things, but because they care

non-derivatively for doing what God commands. Thus, it looks like we have intuitions on both sides of the fence here.

Smith's account of the dispositions a good person has is also flawed in other respects. Imagine the case of Sue, who cares about justice, and so cares only that people get equal concern and respect. Due to her direct concern for justice, Sue ensures that people are given a basic minimum income, food, shelter, and so on. But Sue does not care directly about ensuring that people have adequate income, housing and food. Sue cares directly about people being treated justly, i.e., with equal concern and respect. On Smith's view, however, Sue would not count as a good person; her dispositions are not that of a good person's. But this begins to seem counter-intuitive given that we might well describe a person with such dispositions as a good person.

These cases illustrate that our evaluations of individuals usually comes down to the way in which we intuitively conceive of a good and strong-willed person. But if there are intuitions on both sides of the fence, then we do not have a decisive reason to reject the externalist's explanation of the reliable connection between moral judgment and motivation. Both Smith and the externalist face the same predicament: we hit bedrock regarding who is right with respect to who counts as a good person. The only way that Smith can prove that internalism is preferable to externalism would be to provide us with an independent argument for his account of what a good person is. He has failed to do this. Therefore, we have no reason to accept internalism over externalism

But his would be too quick. I cannot help think that there is something not quite right about the people in the examples I have given above. It does appear that a good

person cares for the things she thinks morally important directly. Further, despite the fact that this is just an intuition, it does appear to be both a very strong one and a very good one. Let us see if there is any way in which externalism can take account of it.

Smith suggests that if the externalist wants to explain the reliable connection he has to ascribe to the agent in question a non-derivative desire to do what is right. The other desires which an agent may have are derived from this desire. But the externalist might just reject this proposal. On the altered externalist view, what distinguishes good people from others is that they have many different non-derivative desires, or direct moral concerns, for their family, for the well-being of others or for justice. What explains the reliable connection between motivation and moral judgment on the externalist view is an agent's desire to do what she is obligated to do for its own sake. Let us suppose that I become convinced that something which I now hold to be morally worthy is no longer morally worthy. If I desire to do all those things which I deem to be morally worthy directly, then presumably if I am a good and strong-will person, I will now, given my desire, come to have direct concern for what I now believe is morally worthy. What explains the reliable connection is that I have a non-derivative desire to do what is right for its own sake, and this takes account of the intuitions while leaving the plausibility of externalism intact.

Moreover, Alexander Miller[158] argues that Smith's explanation of the reliable connection between a moral judgement and motivation in a good and strong-willed person "is no explanation at all."[159] Both the internalist and the externalist are attempting

---

[158] Miller, 1996.
[159] Miller, 1996, 171.

to explain the fact that if an agent judges that it is right to Φ in circumstances C then, if the agent is not suffering from some psychological failing, she will be motivated to Φ. But what is *the nature* of this reliable connection between moral judgement and motivation?

Miller suggests that the nature of the connection is either conceptual or empirical. If the reliable connection is conceptual in nature, then Smith has assumed the truth of internalism and so begged the question. If the connection is conceptual in nature, then it is a fact that it is a conceptual truth that if an agent judges it right to Φ, then she will be motivated to Φ, if she is not suffering from a psychological failing. But, as Miller explains: "if we view the reliable connection between moral judgement and motivation in a good and strong-willed person as conceptual in nature, then internalism cannot explain this fact, because the fact to be explained is simply the truth of internalism. The attempted explanation reduces to: *internalism is true because internalism is true*."[160] Thus, since Smith's argument is clearly question begging, it cannot provide us with a more plausible explanation of the reliable connection than externalism.

Now suppose that the reliable connection between moral judgement and motivation is empirical in nature. If so, then it is an empirical truth that if an agent judges that it is right to Φ in C then, if an agent is not suffering from a psychological failing, then she will be motivated to Φ. Miller argues that if this is the case, "the internalist cannot provide an explanation of this fact because it is inconsistent with the statement of internalism itself: internalism claims that this fact obtains as a matter of conceptual

---

[160] Miller 1996, 172.

necessity and so is constrained to deny that it consists in the obtaining of a mere empirical regularity."[161]

In sum: Smith is caught in a dilemma. If the reliable connection between moral judgement and motivation in a good and strong-willed person is conceptual in nature, then Smith's employment of the practicality requirement assumes the truth of the practicality requirement and so begs the question against the externalist. If the reliable connection is empirical in nature, then employing the practicality requirement to explain this fact amounts to a denial of the practicality requirement because it says that the connection is conceptual. Either way Smith's employment of the practicality requirement to explain the reliable connection fails.

In response, Smith argues that it is true that the reliable connection that he is trying to explain is in fact conceptual, but that he does not beg the question.[162] According to Smith, the practicality requirement expresses a conceptual truth. The argument in favour of the practicality requirement begins with the claim that it is a "striking fact" that when a good and strong willed person agent judges it right to Φ, she is motivated to Φ. Both internalism and externalism should explain this fact. However, I have already argued that externalism (in addition to internalism) can explain this fact plausibly.

Smith restates his argument in the following way so as to demonstrate that it does not beg the question against the externalist. The externalist denies that the practicality is a conceptual truth in virtue of the fact that, according to the externalist, it is possible that a person - an amoralist, for instance - can believe that something is right, not be suffering

---

[161] Miller 1996, 173.
[162] Smith 1996, 175-176.

from a psychological failing, but fail to be motivated by that consideration. The externalist thus accepts the following: "If an agent judges it right to Φ in C, and she is an amoralist, then, even if she does not suffer from weakness of will and the like, she will not be motivated to Φ in C."[163] Smith calls this Amoralist Externalism (AE). If it turns out that this statement is true, then the practicality requirement is false. Smith argues that it is impossible for there to be an amoralist and, hence, that the practicality requirement is not false.

Smith suggests that if the externalist accepts the above conceptual truth about the amoralist, then he is committed to accepting the following conceptual truth regarding the moralist: "If an agent judges it right to Φ in C, and that agent in a moralist, then she is motivated to Φ in C, at least absent weakness of will."[164] Smith calls this Weak Moralist Internalism (MI). What marks the difference between the moralist and the amoralist for the externalist is that moralists possesses the following "executive virtue": "the virtue of being disposed to conform their motivations to their moral beliefs in a reliable way, at least absent weakness of will and the like."[165] Smith maintains that MI is a statement of the "striking fact" that constitutes the premise in the argument in favour of the practicality requirement. The striking fact is this: if an agent judges it right to Φ in C, and that agent is good and strong-willed, then she is motivated to Φ in C. If this is the case then Smith's argument no longer begs the question, since the externalist is committed to holding it as a conceptual truth. As Smith explains: "Just as Weak Moralist Internalism is a conceptual

---

[163] Smith 1996, 176.
[164] Smith 1996, 176.
[165] Smith 1996, 177.

truth, a conceptual truth to which externalists are committed by their definition of the 'amoralist', it follows that the striking fact is a conceptual truth as well - indeed the very same conceptual truth - and that externalists are therefore committed to it."[166]

Smith suggests that since both internalist and externalist alike must accept MI, they both must give an explanation of it. Smith argues that only the internalist can explain MI in a plausible way. But if the argument I gave above is right, then Smith is wrong. Moreover, is it true that the externalist must accept MI ? Smith argues that the externalist is committed to embracing MI because he holds AE. But the externalist might well jettison the claim about the possibility of the amoralist, but still maintain that the connection between moral judgement and motivation is purely contingent, which entails rejecting MI. That is, the externalist can reject the claim that he holds AE and so reject that he holds MI. Rejecting the claim that there are amoralists does not amount to letting go of one of the central claims of externalism. After all, the main impetus behind embracing externalism follows from the fact that an externalists embrace the Humean account of motivation. On this view, what motivates an agent to act is not the moral belief on its own, but the moral belief in conjunction with a desire (which is external to the moral judgement). Now, Smith accepts that the Humean account of motivation is right, so he cannot object to the externalist account of motivation. And, if the externalist drops his account of the amoralist, then he is no longer committed to accepting MI as a conceptual truth, given that he was forced to do this because he held AE. But MI represents the first premise in Smith's argument in favour of the practicality requirement.

---

[166] Smith 1996, 177.

If the externalist rejects the first premise in the argument as false (given that Smith argued that the first premise was something that both the internalist and the externalist accept), then the argument in favour of the practicality requirement is unsound. If Smith relies on the claim that the MI has to be accepted by both the internalist and the externalist in order to refute externalist, and the externalist can deny that he holds MI, then Smith has no argument in favour of the practicality argument or against externalism.

4.2 Rationalism

The second form of internalism which Smith defends is called rationalism. Rationalism is the claim that if an agent believes that she ought to Φ in C, then there is a reason for her to Φ in C. The main point behind Smith's rationalism is as follows: "It is a platitude that an agent has a reason to act in a certain way just in case she would be motivated to act in that way if she were rational. And it is a consequence of this platitude that an agent who judges herself to have a reason to act in a certain way - who judges that she would be so motivated if she were rational - is practically irrational if she is not motivated to act accordingly."[167]

Philippa Foot[168] argues that we should reject the rationalist claim that moral obligations, which express facts, are facts about our reasons for action. Foot's argument begins by noting Kant's distinction between categorical and hypothetical imperatives. Hypothetical imperatives express reasons for action just if the agent to whom the imperative is directed has a desire which is served by acting on the imperative. Absent the desire, the imperative no longer gives the agent reason for action. In contrast, categorical

---

[167] Smith 1994, 62.
[168] Foot 1972.

imperatives are requirements which give agents reason for action whether or not the agent to whom the imperative is directed has a desire to act on the imperative or has a desire which is served by acting on the imperative. Kant thinks that the requirements of morality are categorical imperatives and, thus, are requirements which give agents reason for action independent of desires. Roughly speaking, Kant holds what Smith calls rationalism.

Foot agrees with Kant that moral requirements express categorical imperatives. They are categorical imperatives because the 'should' claim in a moral judgement - for example, 'You should not murder' - is buttressed by pointing out some feature of the agent's situation, a situation which other agents face with different desires and circumstances. Foot suggests that if we accept this fact, we have to recognize that the very same categorical use of 'should' is found in expressions of the requirement of etiquette. For instance, suppose I contravene a requirement of etiquette by belching at the dinner table. Foot argues that it is right to say of me that I should not have acted as I did despite the fact that acting in accord with the requirements of etiquette serves no desire of mine. As with the requirements of rationality, the 'should' claim of etiquette is buttressed by pointing out some salient feature of my circumstances, circumstances which are faced by other agents with different desires, and do on.

Foot maintains that by reflection on the fact that the requirements of etiquette and the requirements of morality are both categorical, we see that the exponent of the rationalist conceptual claim is caught in a dilemma. No one puts forth the claim that requirements of etiquette are requirements of reason or rationality despite the fact that

they express categorical imperatives. One who disobeys rules of etiquette is not thereby irrational. And, since the rationalist will be forced to admit this, he is forced to admit that the requirements of morality are not requirements of reason or rationality.

If the rationalist wants to hold that the requirements of morality are requirements of reason, he will have to provide an argument to support this claim. But Foot argues that to hold this view would be to say something false, since there is nothing irrational in asking whether one has reason to act in the way morality enjoins. This follows from the fact that Foot thinks that irrational actions are "those in which a man in some way defeats his own purposes, doing what is calculated to be disadvantageous or to frustrate [an agent's] ends."[169] Failing to act on moral judgements results in no such irrationality. Thus, the requirements of morality are not requirements of reason and, hence, rationalism is false.

Smith is not persuaded by Foot's view according to which morality is a system of hypothetical imperatives.[170] Moreover, Smith suggests that "there is a single, powerful line of argument in favour of the rationalist conceptual claim. The argument trades on the truism that we expect agents to do what they are morally required to do."[171] Smith's "powerful" argument can be put as follows: Moral requirements apply to agents insofar as they are rational. And, it is a conceptual truth that if an agent believes that he is under a moral obligation, and he is rational, we expect him to act in the way his moral obligation enjoins. Being rational "must therefore suffice to ground our expectation that rational

---

[169] Foot 1972, 310.
[170] Smith 1994, 80-85.
[171] Smith 1994, 85.

agents will do what they are morally required to do."[172] This can be the case, Smith argues, only if we think that "moral requirements that apply to agents are categorical requirements of reason."[173]

To Smith's mind, the crucial premise in this argument is that we expect rational agents will do as their moral beliefs enjoin. How does Smith justify this crucial premise? Smith argues that what justifies this premise is that "the appropriateness of a whole range of moral attitudes depends on the truth of the rationalist's conceptual claims," i.e., the crucial premise. It is a truism that we regularly disapprove or approve of people's behaviour from the moral point of view. When an agent acts wrongly we disapprove of her action, and when an agent acts rightly we approve. And the attitudes of approval and disapproval are apposite only "when there exists *grounds* for legitimate expectation about how someone will behave."[174] We can disapprove of an agent only if we expect rational agents to do what they morally ought to do. The legitimacy of this expectation is grounded in the fact that people are rational. As Smith puts it: "Being rational suffices to ground the expectation that people do what they are morally required to do. Given that moral approval and disapproval are ubiquitous, the truth of the rationalist conceptual claim seems to be entailed by the fact that the preconditions of moral disapproval and approval are satisfied."[175]

Unfortunately, Smith's claim that the fact that it is legitimate to approve or disapprove of agent's actions presupposes that it is plausible to assume that people will

---

[172] Smith 1994, 85,
[173] Smith 1994, 85.
[174] Smith 1994 89.
[175] Smith 1994, 90.

act morally. To my mind, this claim seems radically false. Consider the following thought experiment.[176] Imagine that we come in contact with a world, call it w_1, which is similar to our own world, w_0, with the exception that the inhabitants of world w_1 are consequentialists while we are deontologists. Thus, the moral theory of the inhabitants of w_1 picks out the natural property Z as the property in virtue of which things are right. But the inhabitants of w_0 are deontologists and their moral theory picks out the property Y as the property in virtue of which things are right. Suppose we invite an inhabitant of w_1 to our world, w_0. Upon arrival, the inhabitant of w_1 begins to commit all kinds of acts which, according to the moral views of w_0, are immoral. Now, we know that we cannot expect the inhabitant of w_1 to act in the way we think that agents of w_0 morally ought to act. Yet, it seems to be the case that to disapprove of his actions is, in a strong sense, legitimate. But if it is legitimate to disapprove of his behaviour without presupposing that it is reasonable to assume that he will act morally (as per w_0), then it follows that it is simply not true that our practice of approval and disapproval at w_0 presupposes that we believe that agents from any world, w_0, w_1, w_2, . . . w_n, will act morally (as per w_0). Suppose, further, that the newly arrived guest begins to observe our practices and finds them utterly reprehensible. Might he also not cry out "I disapprove of what you w_0 dwellers are doing." If so, then it again follows that it is not a precondition of disapproval and approval that we assume that agents will act as morality enjoins at w_0.

Smith might dissent to this story on the grounds that the inhabitants of w_1 did not understand that what they were doing was reprehensible to the inhabitants of w_0. Perhaps

---

[176] For the purposes of this thought experiment, I will assume that the inhabitants of the different worlds have equally justified moral views.

once they arrived at knowledge of the moral conventions that we have, they might have been expected to act morally. But this reply will not help Smith because the visitors might have learned all there is to know about our moral conventions, but still not be inclined to act in the way we think they morally ought. Yet, we would still think that it was appropriate to disapprove of their behaviour despite the fact that we have no such belief regarding the fact that agents typically act as morality enjoins at w_0.

It is no matter that my objection is merely the product of philosophical fantasy. Let us consider a less fanciful example. Suppose that there are two families living side by side. One family embraces consequentialism and one family espouses deontology. Now, it is clear that the two families are going to have divergent moral views. Further, given that they are familiar with each other they both know what each other's moral views will be given a certain situation. Now, both families know that the other will act contrary to each other's moral views. However, this does not stop the families from disapproving of each other's behaviour. In fact, we need no such presupposition to ground our practice of disapproval and approval. And, Smith's argument in favour of rationalism requires that the premise, according to which we expect rational agents will do as their moral beliefs enjoin, be true. Since it is false, we have no reason to accept his view of rationalism.[177]

4.3 Smith's Account of Normative Reasons

In this section I will outline Smith's argument in favour of the claim that moral judgements are both objective and practical. Smith attempts to demonstrate this claim by

[177]Moreover, perhaps our practice of approval and disapproval presupposes to a certain extent that people will not act as they morally ought, and that this is why that practice is so widespread. Why else would we bother to approve and disapprove if we didn't think that this could make people do what they otherwise would not.

distinguishing between what he calls motivational reasons and normative reasons. Consider the claim that "F has a reason to Φ." This claim can be understood in two ways. Smith suggests that to say that F has a normative reason to Φ is equivalent to saying "that there is some normative requirement that [F Φs] and is thus to say that [F's Φ-ing] is justified from the perspective of the normative reason that generates that requirement."[178] According to Smith, normative reasons are best thought of as truths or propositions about what we ought to do. The salient feature of normative reasons is that they justify action. Motivational reasons, on the other hand, are not justificatory. As Smith explains: "The distinctive feature of a motivating reason to Φ is that, in virtue of having such a reason, an agent is in a state that is *explanatory* of her Φ-ing, at least other things being equal."[179] Since an agent who has motivating reasons is in a state which is potentially explanatory of her action, motivating reasons are better thought of as psychological states. With this distinction in mind, we can turn to Smith's analysis of normative reasons which, he contends, is the key to solving the moral problem.

On Hume's account of motivational reasons, the reasons which motivate agents are comprised of both a desire and a means-end belief. Hume also thought that what justifies an agent in acting is that this would to make it as likely as possible that she get what she most wants. That is, an agent acts rationally when she maximizes on her currently held desires. This is Hume's account of rationality and, thus, his account of normative reasons. Smith argues that Hume's account of motivational reasons is correct,

---

[178] Smith 1994, 95.
[179] Smith 1994, 96.

but that this account of normative reasons, Hume's account of rationality, should be rejected.

Smith's account of normative reasons begins with a puzzle. Recall the distinction between motivating and normative reasons. Smith argues that we can employ both types of reasons to explain action and they are thus in potential conflict with each other. According to Smith, there are two perspectives from which to explain human action: the intentional and the deliberative. From the intentional perspective we explain action as resulting from an agent's beliefs and desires. Her action is caused by her motivating reasons. Yet, we can also explain action from the deliberative perspective. When explaining action from this perspective we do not cite an agent's motivating reasons to explain her action; we explain an agent's action by appeal to her values, i.e., what she takes to be her normative reasons. The deliberative perspective is a perspective from which to explain action because it appears that deliberating on what to do given her values an agent can be moved to act. Smith argues that the problem here "is to explain how deliberation on the basis of our values can be practical in its issue *to just the extent that it is*."[180]

Agents deliberate; as a consequence of deliberation on what to do given her values, an agent can be moved to act. In order to explain just how this happens it is important to figure out just what it means to value something. Smith suggests that there are two candidates: valuing can be understood in terms of belief, or in terms of desire. Smith rejects the claim that valuing amounts simply to desiring.[181] Thus, he is left with

---

[180] Smith 1994, 136.
[181] Smith 1994, 137-147.

the claim that valuing is believing. But if valuing is believing, and if Hume is correct that an agent is motivated just if she has a desire (and a means-end belief), how can valuing be both a belief and be practical?

Smith argues that there is a conceptual connection between believing that something is desirable and desiring. The conceptual connection can be expressed as follows: "If an agent believes that she has a normative reason to Φ, then she rationally should desire to Φ."[182] This conceptual truth follows from Smith's analysis of normative reasons. Two questions arise with respect to the conceptual connection between believing something desirable and desiring. First, what is it for an agent to have a normative reason? Second, what is it to be rational?

To answer the first question Smith attempts to give an analysis of normative reasons which demonstrate the connection between believing something desirable and desiring. According to Smith, "it is a platitude to say that it is desirable that we do what we would desire to do if we were fully rational; that what we have normative reason to do is what we would desire that we do if we were fully rational."[183] Smith's analysis of this platitude turns out to be the following:

> The platitude tells us that what it is desirable for us to do is what we would desire that we do if we were fully rational. In other words - and now we are turning the platitude into an analysis, for we are making explicit distinctions that are at best only implicit in the platitude - it tells us that what it is desirable to do in certain circumstances - let's call this the 'evaluated possible world' - is what we, not as we actually are, but as we would be in a possible world in which we are fully rational - let's call this the 'evaluating possible world' - would want ourselves to do in those circumstances. That is, it tells us that facts about the desirability of acting

[182] Smith 1994, 148.
[183] Smith 1994, 150.

in certain ways in the evaluated world are constituted by facts about the desires we have about the evaluated world in the evaluating world."[184]

In this way, Smith demonstrates that moral judgements are beliefs about what an agent has normative reasons to do, and what he has normative reasons to do is what he would desire to do if he were fully rational. And, if we believe that an agent would desire something if rational, then the agent ought rationally to desire the thing in question. Moral judgements thus express beliefs about the reasons agents have for action.

Smith still needs to answer the second question regarding the nature of a rational agent. Bernard Williams[185] argues that to be fully rational requires that an agent have no false beliefs, all relevant true beliefs, be instrumentally rational and have a vivid imagination. Smith adds to this list a reflective equilibrium condition. According to Smith, fully rational deliberating agents bring their desires into a sort of systematic unity by adding more specific or more general desires to their set. This last condition is important for Smith because he thinks that on the basis of proper deliberation (e.g., in Williams's case exercise of the imagination, and in Smith's case the exercise of systematic justification) an agent can arrive at new, underived desires and destroy old underived desires.

Thus far we have not really said how it is that Smith's account of normative reasons is anti-Humean. The distinction between Smith and the Humean here lies in the way in which they conceive of normative reasons. Smith believes that the judgement that there is a normative reason for agents to act is equivalent to the judgement that it is

---

[184] Smith 1994, 151.
[185] Williams 1979.

desirable that they act. Hume thought that normative reasons were relative to agents, given their different motivational sets. Smith takes the opposite route, arguing that normative reasons are non-relative: thus, if an agent has a reason to Φ in circumstances C, then everyone has a reason to Φ in C. Smith puts the point this way: "Consider, for example, the schematic claim 'It is desirable that p in circumstances C.' On the non-relative conception of normative reasons, this claim has a straightforward truth condition: it is desirable that p in C just in case we would all desire that p in C if we were fully rational."[186] Here is where Smith's account of rationality differs sharply from Hume's. On Hume's view the concept of practical reason is inherently relative: what x has reason to do depends on the desires of x and what y has reason to do depends on the desires that y has, and so on. In the following section, I will critically evaluate Smith's account of normative reasons.

4.4 Smith's Account of Normative Reasons Undermined

To recap, on Smith's view, moral beliefs are beliefs about what an agent thinks she has normative reason to do, and someone believes that she has a normative reason to carry out some behaviour, only if she believes that she would desire to carry out that behaviour, if she were rational. We can link this last claim with motivation by demonstrating the conceptual connection between an agent's normative reasons and her desires in the following manner: if an agent believes that were she rational, she would desire some course of action, then she is rational just if she desires to carry out that course of action. Roughly speaking, the reasons which one has for action depend on the truth of the

[186] Smith 1994, 166.

counterfactual: an agent desires to do whatever she would want herself to do if she were rational, where being rational satisfies Smith's constraints on rationality.

The viability of Smith's link between what an agent has normative reason to do and what she desires to do turns on the claim that if an agent believes that were she rational, she would come to desire Φ, then she is rational just if she desires to Φ. Elsewhere Smith says that the belief that an agent would desire something if she were rational causes in the rational agent a desire to do what she believes she has normative reason to do. But why should this (a) cause in an agent a desire and (b) be accepted by the Humean? It remains utterly mysterious how it is that the mere fact that an agent believes that she would desire something were she rational, causes her to have a desire and be motivated. For, as Hume thought, beliefs and desires are distinct existences. But if this is the case, then it does not appear to be the case that Smith has established that a belief can cause in an agent a desire to do what beliefs enjoin. Of course, Smith might argue that an agent might have a moral desire, e.g., a desire to act morally. This moral desire could explain why the belief that there is a normative reason in favour of a certain course of action causes in agents a desire to act morally; the desire is acquired as a means to satisfying the moral desire, i.e., the desire to act morally. Obviously, however, Smith would reject this possibility: for this explanation flies in the face of his claim that it is the content of a moral judgement which gives agents reason for action insofar as they are rational. But then it is still not clear how the belief causes in agents a desire to do what their belief enjoins. Moreover, why should the Humean accept that an agent has reason to do an action on the grounds that, in a counterfactual situation in which she was fully

rational, she would desire to do that action? For Hume, the counterfactual claim represents a mere belief, and beliefs do not motivate, so why should this belief motivate? Smith seems to have failed to defend this central aspect of his account of normative reasons.

Furthermore, Christian Piller[187] argues that the counterfactual account of practical reason faces a dilemma. As Piller explains: "either it mistakes successful action for rational action (if what we have reason to do is identified with what we would desire to do if we were fully rational), or it imposes either a mistaken or a useless standard of rationality (if what we have reason to do is identified with what we *believe* that we would desire to do if we were fully informed)."[188] Smith suggests that the central tenet behind his account of normative reasons is the claim that an agent has a reason to do what she would desire to do were she fully rational.

The first horn of Piller's dilemma is as follows. Smith thinks that normative reasons are the reasons which justify our actions. Now consider the following situation put forth by Smith:

> Suppose I now desire to drink a gin and tonic and believe I can do so by mixing the stuff before me with the tonic and drinking it . . . Suppose further that this belief is false; that the stuff before me is petrol rather than gin. Surely it would be appropriate for the outsider to say I have no reason to mix this stuff with tonic and drink it. Yet I have both the relevant belief and desire.[189]

---

[187] Piller 1996.
[188] Piller 1996, 361.
[189] Smith 1994, 94; See also Williams 1979, 18.

For the sake of argument let us suppose that I do not mix the tonic with the petrol. According to Piller, Smith seems committed to the admission that since I act in accord with my normative reasons I thereby act rationally. Yet, upon reflection, it seems implausible to assume that I did act rationally; it was solely a matter of luck that I acted in accord with my normative reasons given that I was ignorant of them. Piller puts his objection more generally as follows: "we act in accordance with our normative reasons if we act in the way that turns out to be best for us. But acting rationally and acting in a way that turns out to be best for us are not the same. Smith's theory of normative reasons should not exclude the possibility that rational actions, i.e., actions justified by normative reasons, might fail to achieve their ends. To capture this conceptual possibility is a condition of adequacy on any theory of rationality."[190]

Now, Smith might attempt to circumvent what look to be counter-intuitive implications of his theory of rationality by suggesting that not only must an agent act in accord with normative reasons she has, but she must also believe that she has these reasons. But this seems to be a costly move for Smith. If it is a precondition that for an agent to be rational she must not only act in a accord with her normative reasons but also believe that she has those normative reasons, then his account of rationality becomes implausibly strong. Here's why. In the example above, whether an agent drinks the petrol and tonic or not, she would have acted irrationally. If she drank the petrol and tonic, she would have acted irrationally because she would have acted against the normative reasons she has. However, if she did not drink the petrol and tonic, then she acted irrationally

---

[190] Piller 1996, 362.

since she did not act in accord with the normative reasons she had because she was unaware that she had them. Thus, either Smith's counterfactual account of practical reasons arrives at the wrong results or it is useless (on the grounds that it fails to render any results at all).

Smith also suggests, contra the Humean, that the (normative) reasons we have are requirements of rationality and these reasons are not agent-relative. The concept of having a reason for action is a non-relative concept in the sense that: if an agent has a reasons to Φ in circumstances C, then every rational agent who faced the same circumstances has reasons to Φ. This is how Smith establishes that the requirements of morality are categorical imperatives, i.e., requirements of reason which bind on all agents.

But I do not think that Smith has established the fact that reasons for action are non-relative. To illustrate, consider the other-worldly types which we saw in 4.2. The difference between the visitors from the other planet, inhabitants of w_1 and the inhabitants of our world, w_0, is that they are consequentialists and we are deontologists. Thus, they differ over what nature property an action has in virtue of which it is right or wrong; the inhabitants of w_1 think that it is natural property Z, while the inhabitants of w_0 think it is natural property Y. Now, suppose that the inhabitants of w_0 have satisfied Smith's requirements of rationality, i.e., we have no false beliefs, all relevantly true beliefs, we have used our imagination and our desires are systematically justified in reflective equilibrium. As a consequence of satisfying these requirements we have converged on what it is rational to do in certain moral situations. Now, suppose that the inhabitants of

w_1 have read Smith's book and have attempted to satisfy Smith's requirements of rationality. Imagine that they have no false beliefs, all relevantly true beliefs, and they have deliberated so that their desires are systematically justified in reflective equilibrium. As a consequence of their satisfying Smith's requirements, they too have converged on what it is rational for agents to do in certain moral situations. Yet each of the two groups has completely different sets of moral reasons.

Let us suppose that we all get together at an inter-galactic meeting, and at the meeting we witness a certain act. In response, the humans react with horror, but the w_1 inhabitants seem unmoved; in fact, they seem glad about it. Now each group attempts to convince the other of its moral position. However, since each groups begins with a different starting point - for example, the members of the group differ in their most basic desires - they cannot begin to agree on anything and thus remain unmoved by the other groups.

The obvious response here is to argue that either the w_0 dwellers or the w_1 dwellers have not reached full rationality. But that does not seem to be the case. The problem with Smith's account of rationality is that it imposes no constraints on the first values or desires that an individual, or group of individuals, can have. As a result of this, people with radically different intuitions, desires, values, and so on, can arrive at radically different, yet fully rational, accounts of what they have normative reason to do. Smith requires a constraint on what it is to have a rational first value; but he has failed to provide such a constraint. Smith might just argue that the scenario which I have constructed is just the product of philosophical fantasy, and not a realistic possibility. But

this would be a mistake, given that, *mutatis mutandis,* we could change the two groups from consisting of people who are from other planets to people who are from different societies. It is at least possible that people who come from different societies will arrive at radically different moral judgements, though they have satisfied the requirements of full rationality. It is in virtue of the fact agents may begin with any desires whatever that agents arrive at different fully justified accounts of normative reasons. Thus, Smith's account of normative reasons is not agent-neutral. Since Smith requires an agent-neutral account of normative reasons to conclude that moral judgments can be objective and practical, and he fails to provide one, Smith's answer to the moral problem fails.

4.5 Conclusion

In this chapter, I have argued that Michael Smith has failed to establish the truth of internalism, and that his account of normative reasons is flawed. Smith has thus failed to solve the moral problem. Like David McNaughton and Jonathan Dancy, Smith cannot demonstrate that moral judgements are inherently practical. For the most part, I have not defended any positive view, though I have attempted to provide at least a *prima facie* case in favour of externalism about moral motivation; according to this case, an agent is motivated not by a moral judgement itself, but by things external to that judgement, e.g., the agent's desires. Moreover, I have attempted to defend the Humean account of motivation against various objections. In conclusion, I think that the most plausible solution to the moral problem requires rejecting the claim that morality is inherently practical.

BIBLIOGRAPHY

Altham, J. E. J. 1986. "The Legacy of Emotivism", Fact, Science and Morality: Essay on A. J. Ayer's Language, Truth and Logic. Eds. Graham Macdonald and Crispin Wright. Oxford: Blackwell. 275-88.

Ayer, A. J. 1936. Language, Truth and Logic. London: Gollancz.

Brink, David. 1989. Moral Realism and the Foundations of Ethics. Cambridge: Cambridge University Press, 1987.

Dancy, Jonathan. 1993. Moral Reasons. Oxford: Basil Blackwell.

Darwell, Stephen, Allan Gibbard, and Peter Railton. 1992. "Toward *Fin de Siecle* Ethics: Some Trends." The Philosophical Review 115-190.

Foot, Philippa. 1972. "Morality as a System of Hypothetical Imperatives", Philosophical Review. 305-316.

Gauthier, David. 1975. "Reason and Maximization", Canadian Journal of Philosophy. 411-34.

Hare, R. M. 1981. Moral Thinking. Oxford: University Press.

Harman, Gilbert. 1977. The Nature of Morality. Oxford: Oxford University Press.

Hume, David. 1888. A Treatise of Human Nature. Ed. L.A. Selby-Bigge. Oxford: Clarendon Press, 1978.

Kant, I. 1786. Foundations of the Metaphysics of Morals. Library of Liberal Arts, 1959.

Mackie, J. L. 1977. Ethics: Inventing Right and Wrong. London: Penguin.

McNaughton, David. 1988. Moral Vision: An Introduction to Ethics. Oxford: Blackwell.

Miller, Alexander. "An Objection to Smith's Argument for Internalism", Analysis. 169-174.

Nagel, Thomas. 1970. The Possibility of Altruism. Oxford: Clarendon Press.

Piller, Christian. 1996. "Critical Notice of Michael Smith, The Moral Problem", Australasian Journal of Philosophy. 347-367.

Platts, Mark. 1979. Ways of Meaning. London: Routledge and Kegan Paul.

Smith, Michael. 1987. "The Humean Theory of Motivation", Mind. 36-61.

_____. 1991. "Realism." A Companion to Ethics. Ed. Peter Singer. Oxford: Basil Blackwell. 399-410.

_____. 1994. The Moral Problem. Oxford: Blackwell.

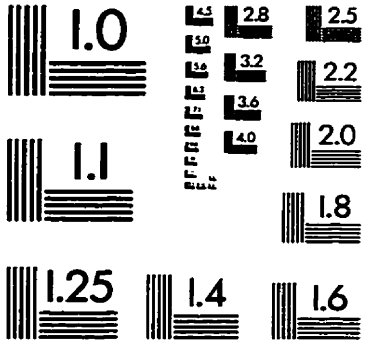
_____. 1996. "The Argument for Internalism: Reply to Miller", Analysis. 175-184

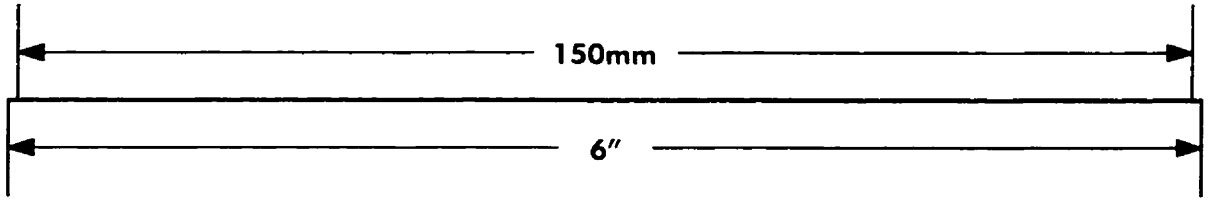
Stevenson, C. L. 1937. "The Emotive Meaning of Ethical Terms" Mind. 14-31.

Stocker, Michael. 1979. "Desiring the Bad: An Essay in Moral Psychology", Journal of Philosophy. 738-53.

Watson, Gary. 1987. "Responsibility and the Limits of Evil. Responsibility, Character and the Emotions: New Essays in Moral Psychology. Ed. Ferdinand Schoeman. Cambridge: University Press. 256-286.

Williams, Bernard. 1979. "Internal and External Reasons?" Rational Action: Studies in Philosophy and Social Science. Ed. Ross Harrison. Cambridge: Cambridge University Press. 17-28.





APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989