THE MOTIVATING FORCE OF MORAL BELIEFS

A thesis submitted in partial fulfilment of the
requirements for the degree of

MASTER OF ARTS

In

THE DEPARTMENT OF PHILOSOPHY

of

RHODES UNIVERSITY

By

James Antony Vos

February 2005

Abstract

I investigate the issue of whether or not one's holding a moral belief is sufficient to motivate one to act as that belief prescribes. I argue that rational persons who hold a moral belief that is also a 'self-referential belief' will form a desire to act as that belief prescribes and thereby be motivated to act on the moral belief.

I argue for this claim by, firstly, showing that the demand that moral judgements must be intrinsically motivating, Internalism, should only apply to rational persons, that is, the link between moral judgement and motivation can be broken in cases of irrationality.

Secondly, I argue against the Humean claim that one cannot rationally form a desire simply because one believes that one ought to have that desire. This claim requires an investigation into a variety of views of Practical Reason and an argument concluding that Practical Reason is broader in scope than the Instrumentalist or Humean allows. I undertake this task in chapter 2.

Thirdly, I argue that believing that I ought to perform a certain action will give me an internal commitment to perform that action, insofar as I am rational. I argue that an internal commitment is a form of desire. Once I recognise that I have a moral belief and an internal commitment, I will be motivated to act as that belief and commitment prescribe.

Acknowledgements

I am indebted to my supervisor, Professor Marius Vermaak, for assisting me greatly in the planning and writing of this thesis.

I must thank my classmates and close friends, Christopher Wareham and Gareth Watkins, for their support and advice over the past year. I am grateful for the discussion and input from a number of other current or past Rhodes University postgraduates: Richard Flockemann, Michaela Baker, Eusebius Mackaiser, Veli Mitova, Nicholas Ferreira, and Nikolai Viedge.

I have enjoyed the support I have received from the Rhodes University Philosophy Department staff. In particular I must thank Dr Samantha Vice, Dr Ward Jones and Francis Williamson. The caring and help I have received from the department secretary, Mrs Merle Murray, has been wonderful.

Thank you to my housemates from 2004 for keeping me cheerful during trying periods of the year. I must thank the Rat & Parrot for that same purpose.

Finally, and most importantly, I must thank my loving parents, Ian and Cilla Vos, for the financial, emotional (and every other type I have wanted) support they have given me.

Contents

Introduction

Determining which actions are morally right and wrong should play an important part in influencing how we act. This, however, is only the case if ethics is practical, that is, if judging that an action is morally right motivates one to act as that belief prescribes. The view that moral beliefs directly motivate action is problematic because of the Classical or Humean theory of Practical Reason, which claims that one requires the presence of a desire and an appropriately related belief in order to be motivated. A belief, on its own, cannot motivate action. How, then, can judging an action to be morally right (where this expresses a belief in the rightness of the action) motivate one to act according to that judgement? Some philosophers – Externalists who, by definition, reject the Internalist claim that moral judgements are intrinsically motivating – attempt to solve this problem by denying that ethics is practical and claim that moral beliefs motivate only if they are accompanied by an appropriate desire such as a second order desire to do what is right. Other philosophers, Non-Cognitivists, Emotivists or Expressivists, deny that moral judgements are forms of beliefs and claim that a moral judgement is more like a desire, emotional reaction or pro-attitude and is therefore intrinsically motivating. Anti-Humean theorists about practical reason deny the claim that belief, on its own, cannot motivate action and develop theories of Practical Reason that aim to show how certain beliefs do motivate action.

The aim of this thesis is to solve this problem in meta-ethics. I intend to reconcile three philosophical theses that are often held to be incompatible. These theses, to state broad definitions, are:

1. Ethical Cognitivism: a moral judgement can be true or false because it expresses, constitutes, or just is, a belief.

2. Internalism: moral judgements are internally and intrinsically motivating. If I judge that eating meat is immoral, then I will be motivated to stop eating meat *because* of my moral judgement. I do not need an extra desire to stop eating meat to bring about the motivation.

3. Humean Theory of Motivation: an agent is motivated to act in a certain way if, and only if, she has an appropriate desire and means-end or constitutive belief, where belief and desire are, in Hume's terms, distinct existences.[1]

My strategy is to reconcile any version of Ethical Cognitivism with a correct interpretation of Internalism and my account of Practical Reason that is Humean in one way and anti-Humean in another. My account is Humean because I insist on the presence of both a belief and desire to motivate action, where belief and desire are defined and correspondingly distinguished according to their different "directions of fit". It is anti-Humean because I allow, and require for my conclusion to follow,

---

[1] Smith, M. *The Moral Problem*. 1994 Blackwell Publishers LTD. p. 12.

that it is possible and rational for one to come to have a desire to, say, refrain from meat eating, because one believes that eating meat is morally wrong.

I begin my argument in chapter 1 by giving an exposition of this debate about moral psychology. I explain each of these theses and discuss the different interpretations thereof. My argument to the conclusion that moral beliefs can motivate compliant action without the presense of an appropriately related desire proceeds as follows:

1. I argue for my interpretation of Internalism, which I name, following Audi, "Rational-Agent-Motivational-Internalism" (RI): If an agent judges that she ought to perform (or refrain from performing) an action, she will be motivated to perform (or refrain from performing) that action because she holds that judgement, and, if she is not so motivated, she is practically irrational.

2. I argue against the Instrumentalist claim that Practical Reason operates only to work out the means to one's ends and does not function to specify what one's ends are or ought to be.

3. I evaluate theories of Practical Reason and argue against Hume and Williams's claim that one cannot arrive at a motivation to Φ (where Φ is some action) through believing that one ought to Φ in a case where one does already, directly or indirectly, desire to Φ.

4. I argue that there must be at least one categorical norm of Practical Rationality because this is necessary for Practical Rationality and reasons for action to exist at all. I argue against Nihilism, which is the view that there is no such thing as Practical Rationality.

5. I argue that there must be some further norms of Practical Rationality, such as to specify what one's ends actually are and to "get the world right" (evaluate one's situation properly and learn what is important from experience), if there is to be at least one norm of Practical Rationality.

6. I evaluate competing theories of desire and its role in Practical Reason and argue that we should understand desire in terms of its direction of fit. It is necessary and sufficient for a state to be a desire that it form a proposition with a specific structure: "A desires that p", where p refers to a state of affairs, for example, that the government ban the selling of meat products. Or "A desires to Φ", where Φ is some action. What these two formulations have in common is that they are directed towards some outcome (an action or state of affairs) that should be the case.

7. I argue that a conception of belief that does not allow one to form a desire to Φ because one believes that one ought to Φ is odd and should be rejected.

8. I discuss a special type of belief through which persons form desires. A "self-referential belief" is a sincerely held belief about what *I* should *do*.

9. I show that self-referential beliefs, *qua* beliefs, have the condition of satisfaction that they be true. Self-referential beliefs are special because they impose another condition of satisfaction on the believer because they are beliefs about what *I* should *do*.

10. Being committed to the truth of one's self-referential belief means that one is committed to believing that one has a commitment to act as that belief prescribes.

11. I recognise the distinction between an external and internal interpretation of "commitment". The external interpretation is that one has a commitment to act that is generated by a normative framework that is relevant to the specific commitment. The internal interpretation of "commitment" is that one has a commitment to act, where a commitment is a form of desire as defined according to its direction of fit. Through holding a self-referential belief that one has a commitment (external sense), one comes to believe that one has a commitment (internal sense).

12. If one has formed one's commitment (internal sense) rationally, one will realise through a process of rational practical deliberation that one does, indeed, have the commitment (internal sense) that one believes that one has. The method of rational practical deliberation is consistent with that for which I argued for in claims 2 - 5.

13. I show that my account of how persons come to form commitments (as a form of desire) and be thereby motivated is consistent with Internalism, as I have defined it, and Ethical Cognitivism.

14. I show how my account of moral motivation explains the Amoralist: someone who believes that eating meat is morally wrong but is in *no way* motivated to stop eating meat. The supposed existence of the Amoralist is thought to refute the view that moral beliefs are intrinsically motivating. I show that this is not the case.

In chapter 1, I argue for claim 1. In chapter 2, I argue for claims 2 to 5. In chapter 3, I argue for claims 6 to10. In chapter 4 I argue further for claim 10 as well as for claims 11 to 14. My argument can be further summarised, with reference to the three theses mentioned earlier, as follows: ethics is practical to the extent that moral judgements motivate persons insofar as they are rational. It is practically rational to form a desire (internal commitment) to Φ because one has a "self-referential" belief that one ought to Φ (has recognised an external commitment to Φ). Rational agents are motivated to fulfil their desires (internal commitments). Therefore one will (if rational) be motivated to act on one's moral belief because that belief will give rise to a commitment, which is a source of motivation. This motivation is internal to the moral judgement because one does not require any desire that is not derived from the moral belief itself in order to be motivated to act.

Chapter 1: Meta-ethics, Internalism and the Humean Challenge

Let us suppose that eating meat is morally wrong and that this is objectively true. There have not, however, been any laws implemented prohibiting the sale and consumption of meat. We, as moral philosophers, see it as our task to persuade people that they ought not to eat meat. How are we to achieve this? In order to effect the desired change in *motivation* to refrain from meat eating it seems necessary, and I shall argue sufficient, for rational persons, to bring about the *belief* that eating meat is morally wrong. Whether this motivation leads to action is a separate matter. There are, I assume, people who do believe that eating meat is morally wrong and who, despite this belief, continue to eat meat. Most (if not all) of these people will be motivated to stop eating meat, but, because of a stronger desire to continue enjoying the taste of meat, they do not bring themselves to become vegetarian. Persons of this sort are not of particular interest in this thesis. Imagine, instead, that Alex believes that eating meat is morally wrong and is in *no way* motivated to become vegetarian. It is cases such as Alex's that are puzzling to moral philosophers and that will be at issue in this thesis.

Alex's case is puzzling because of three widely held intuitions in moral philosophy. I call them "intuitions" because, for now, I only want to bring out the intuitions that lie behind the three corresponding philosophical theses that can all be interpreted in different ways: these intuitions seem not to be able to co-exist in this case. The first intuition[2] is that when one judges that an action is morally wrong, one is expressing a belief about what is objectively the case. Alice finds a lost wallet and judges that she ought not to keep it for herself and ought to return it. Ethical Cognitivism holds that Alice's judgement can be true or false – it is more like[3] a belief than a desire. Non-Cognitivism holds that Alice's judgement cannot be true or false, but rather appropriate or inappropriate, or desirable or undesirable – it is more like a desire or attitude than a belief. One could, of course, claim that to hold a moral judgement is not to be in a purely belief-like state or a purely desire-like state, but involves some combination of the two. If moral judgements are non-cognitive – express only a desire-like state and not beliefs at all – they cannot be true or false. The Cognitivist intuition is that a moral judgement is a belief-like state (capable of being true or false) or some combination of a belief-like state and a desire-like state that can be true or false. We (or at least I) would like to think that morality is objective, that certain actions in the relevant circumstances are objectively right or wrong and that this objective status

---

[2] This intuition is not universally shared but it is one central to our ordinary understanding of morality. Nothing in this thesis, however, turns on whether Cognitivist intuitions are universally shared.
[3] I use the deliberately vague phrase "more like" for now, because Cognitivists are not all committed to holding that a moral judgement just is a belief and that is all there is to it.

of morality applies to all persons irrespective of culture and of whatever desires persons contingently happen to have. If moral judgements are non-cognitive, they cannot be objective in this sense.

The second intuition is that it is necessary for the performance of all intentional actions that the actor has a desire and an appropriately related belief (usually means-end). This intuition is expressed in what is commonly called belief/desire psychology.[4] Suppose that I intentionally open a can of Coca-cola and take a sip. I would, in a normal situation,[5] have a desire to drink Coca-cola and a means-end belief that by opening this can of coca-cola and taking a sip I will drink a Coca-cola. Because I am aware of my desire to drink Coca-cola and my belief about how to do it, I will come to have a desire to open the can and take a sip. Both the belief and the desire are necessary, it seems, because without the belief I would not know how to satisfy my desire and without the desire I would be without a source of motivation to drive me to perform the action.

The third intuition is that morality is practical. When one makes a moral judgement, one is motivated to act on that judgement and one does not require an additional source of motivation. The motivation is internal to the holding of the judgement. The philosophical expression of this intuition is Internalism, of which there are various versions. The basic idea behind Internalism is that moral judgements are by their nature intrinsically motivating – the motivation to act on a moral judgement is internal to the judgement itself.

When we combine these intuitions, we get a surprising result. Alex judges that it is morally wrong to eat a plate of spaghetti bolognaise, which is made with meat, because he believes that eating meat is morally wrong. This judgement is a belief-like state that can be true or false. Because Alex holds a moral judgement, we would expect him to be motivated by that judgement to refuse the spaghetti bolognaise and not require any other source of motivation to motivate this refusal because his judgement is itself a source of motivation. In order for Alex to perform an intentional action, in this case refusing the spaghetti bolognaise, he requires the presence of both a belief and a desire, where the belief is needed to tell him how to fulfil his desire. The desire does all the motivating work. But if we combine Cognitivism and Internalism we would expect his judgement (belief-like state) to motivate Alex to act on his judgement without the help of an additional source of motivation, such as a desire. This expectation is precisely what belief/desire psychology denies is possible.

We therefore have a *prima facie* reason to deny Cognitivism or Internalism or belief/desire psychology. If we deny Cognitivism, we deny that one's moral judgements can be true or false and

---

[4] Donald Davidson's view of motivation, "Actions, Reasons and Causes". In *Essays on Actions and Events.* 1980. Clarendon Press. p. 3 – 20, is a paradigmatic example of belief/desire psychology.
[5] We could imagine a situation where a person has been conditioned to open a can of Coca-cola whenever that person feels lust, for example. See Nagel, T. *The Possibility of Altruism.* 1970. Clarendon Press. p. 33 – 34.

therefore that morality is objective in the sense explained earlier. We would also be denying that there are moral facts (that exist or can be known) as denying that moral judgements can be true or false amounts to denying that facts are the appropriate objects of such judgements. If we deny Internalism, we deny that morality is practical, that the holding of moral judgements, itself, motivates action. If we deny belief/desire psychology – that is, deny that having a desire (broadly construed as a motivational state) is necessary for motivation, we seem to be at a loss to explain what does the motivational work in human motivation. Smith claims that the "moral problem" stems from the seeming incompatibility of the following three philosophical theses:

1. Ethical Cognitivism: moral judgements of the form "It is right that I Φ" (where Φ is some action) express a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.

2. The Internalism Requirement: if someone judges that it is right that she Φs then, *ceteris paribus,* she is motivated to Φ.

3. A Humean Theory of Motivation: an agent is motivated to act in a certain way just in case she has an appropriate desire and appropriately related means-end belief.[6]

The three theses are *prima facie* incompatible as together they involve, simply put, that moral judgements express beliefs, all moral judgements are motivating, and belief – on its own – is never motivating. One might wonder if Ethical Cognitivism were interpreted as the thesis that a moral judgement expresses a belief and a desire (where a belief gives rise to a desire), then the incompatibility would be resolved. I believe that this is the kind of move that is required to solve the "moral problem" and the main aim of this thesis is to provide an account of how these three theses can be reconciled. This move is anticipated and blocked by my opponents such as Hume and Williams. Before I can consider their arguments and analyse these three theses in more detail, I must explain where the "moral problem" fits into moral philosophy.

Meta-ethics

The supposition with which I began this chapter, that eating meat is morally wrong, is an example of a normative ethical rule. The claim that Cognitivism is a plausible theory about what it is to hold a moral judgement is an example of a meta-ethical claim. Finding answers to questions such as "Is it morally permissible to eat meat?" and "Should I return the lost wallet that I found?" and justifying these answers within the framework of a set of norms or guidelines is the domain of normative ethics.

---

[6] Smith, M. p. 12.

Meta-ethics asks second order questions about what we are doing when we are debating normative ethical issues such as those mentioned above.[7]

Alexander Miller gives examples of the main sorts of questions that meta-ethics is concerned with:

(a) *Meaning*: what is the *semantic function* of moral discourse? Is the function of moral discourse to state *facts*, or does it have some other non fact-stating role?

(b) *Metaphysics*: do moral facts (or properties) exist? If so, what are they like? Are they identical to some other type of fact (or property) or are they irreducible and *sui generis*?

(c) *Epistemology and justification*: is there such a thing as moral knowledge? How can we know whether our moral judgements are true or false? How can we ever justify our claims to moral knowledge?

(d) *Phenomenology*: how are moral qualities represented in the experience of an agent making a moral judgement? Do they appear to be 'out there' in the world?

(e) *Moral psychology*: what can we say about the motivational state of someone making a moral judgement? What sort of connection is there between making a moral judgement and being motivated to act as that judgement prescribes?

(f) *Objectivity*: can moral judgements really be correct or incorrect? Can we work towards finding the moral truth?[8]

The problem I am dealing with asks the question about moral psychology. This is the question involved in the second of the three seemingly incompatible theses Smith mentions: the Internalism Requirement: If someone judges that it is right that she Φs then, *ceteris paribus,* she is motivated to Φ. I will discuss the various formulations of the Internalism Requirement later in the chapter. The issue of this project is, however, broader than merely finding an answer to the question of moral psychology as stated by Miller. James Rachels calls the problem that this thesis attempts to solve the *motivation problem.* Rachels explains why Internalism is problematic for Cognitivist views of ethics:

… there is an internal connection between moral belief and motivation. Why is this important? For one thing, it poses a problem for the idea that there are moral facts. There is no internal connection between ordinary factual belief and motivation. If I say that Antarctica is bigger than England, this leads to no expectations about how I will behave. It is just a fact, towards which I might be totally indifferent. Of course, if my statement is considered alongside other information about me, such as that I desperately want to visit somewhere bigger than England, then you might conclude I have some motivation to visit Antarctica. But it is the added information about my desires that supplies the motivating power. With moral belief, the motivating power seems built into the belief itself.[9]

This is where the third thesis, the Humean account of motivation, fits in. We want belief in an objective moral truth to have an internal connection to motivation, but no belief is sufficient to motivate. Hume argues from the seeming incompatibility of theses 1 - 3 (Hume accepts 2 and 3) that

---

[7] Miller, A. *An Introduction to Contemporary Metaethics.* 2003. Polity Press. p. 1.
[8] *Ibid.* p. 2.
[9] Rachels, J. *Ethical Theory 1.* 1998. Oxford University Press. p. 3 – 4.

"Morality cannot be founded on Reason alone" - that is, against Cognitivism.[10] There is some dispute as to whether or not Hume was a strict non-Cognitivist in believing that moral judgements do not involve belief at all,[11] but he does think that morality is more "properly felt than judg'd of"[12] and that moral judgements cannot be true or false (Hume means "believed" here when he says "judg'd of"). Hume's claim, as I interpret it, is not that moral judgements do not involve beliefs at all, but rather that a moral judgement cannot be true or false and therefore there are no moral facts. Hume might have held that it can be true or false that a certain moral action or judgement is desirable or undesirable, but that judgement itself cannot be true or false.

Ascertaining the correct interpretation and then truth (or falsity) of Cognitivism, involves answering all of the main meta-ethical questions laid down by Miller. This task is well beyond the scope of this thesis. Philosophers are sceptical about Cognitivism either directly or indirectly. Indirect scepticism involves basing one's *content* scepticism (whether there are such things as moral facts) upon one's *motivational* scepticism (whether these facts, if they exist, can motivate action). Hume's famous argument in *A Treatise of Human Nature* is the prime example: Moral judgements are intrinsically motivating. No belief is ever, on its own, intrinsically motivating. Moral judgements are not, therefore, best thought of as expressing beliefs (alone). Direct scepticism traditionally, though not exclusively, involves what Rachels names the *ontological problem* and the *epistemological problem*.[13] The former problem questions whether moral facts actually exist in the world. Hume writes:

Take any action allow'd to be vicious: Wilful murder, for instance. Examine it in all lights, and see if you can find that matter of fact, or real existence, which you call *vice*. In whichever way you take it, you only find certain passions, motives, volitions, and thoughts. There is no other matter of fact in the case.[14]

The latter, related, problem questions how we can come to know moral facts. There are familiar ways that we can know matters in science, mathematics and in ordinary life through perception. None of these methods is apt for the discovery of moral facts. We are owed an account of how we are to discover these moral facts that exist in the world.[15] John Mackie argues that there are no moral facts that exist in the world by combining these two related objections[16]. Miller summarises his position:

---

[10] Hume, D. *A Treatise of Human Nature.* Selby-Bigge, L.A. [Ed]. 1978. Oxford University Press. Book III. Part 1. Section 1.
[11] Mackie, J. *Hume's Moral Theory.* 1980. Routledge and Kegan Paul. And Snare, F. *Morals, Motivation and Convention: Hume's Influential Doctrines.* 1991. Cambridge University Press. Especially p. 12 -21.
[12] Hume, D. p. 470.
[13] Rachels, J. p. 2 – 3.
[14] Hume, D. p. 468.
[15] Rachels, J. p. 3.
[16] Mackie, J. *Ethics: Inventing Right and Wrong.* 1977. Penguin Books. p. 38 – 42.

…we have no plausible epistemological account of how we could access such [moral] facts and properties, and, moreover, such properties and facts would be *metaphysically queer*, unlike anything else in the universe as we know it.[17]

Giving sound responses to these problems, it should be obvious, is no easy matter. Because of the magnitude of this task (a task not central to my current project), I shall assume that some broad version of Cognitivism is plausible and try to reconcile any version thereof with the best statement of Internalism and a correct account of human motivation. I turn my attention now to analysing the three theses involved in The Moral Problem.

Ethical Cognitivism

Cognitivism in ethics is based on our common sense view that when we judge that something is morally right or wrong, we are holding a belief about an objective matter of fact. There are two broad ways in which we can understand Cognitivism. The first, and weaker, view is epistemic objectivity:

Ethics could be objective in the sense that moral problems can be solved by rational methods. These methods would show that some methods are acceptable while others are not.[18]

An act is right if it is justified by sound reasoning and if it is one that an ideally rational person would assent to. Tests of the moral status of an act (or intention to act) typically include, on this view of objectivity: universalizability, consistency, impartiality, and whether rational persons would accept the act as moral under ideal conditions of deliberation. I see Kant as paradigmatic of this approach. He insists that we should accept as moral only those maxims which we would will to be a universal law of nature. Although this view of morality is objective – there are objective right and wrong acts that we can determine through correct rational reflection – it is also contingent. Morality depends, for its existence, on the presence of rational human minds.

Miller calls this view "Weak Cognitivism":

A *weak Cognitivist* theory is one which holds that moral judgements (a) are apt for evaluation in terms of truth or falsity, but (b) cannot be the upshot of cognitive access to moral properties and states of affairs.[19]

He continues:

This view thus rejects moral realism, not by denying the existence of moral facts, but by denying that those facts are constitutively independent of human opinion.[20]

Epistemic objectivity denies that moral facts are *real* facts or properties that exist in the world. Moral facts, rather, exist as a matter of rational consensus, or potential consensus amongst rational persons in ideal conditions of rational deliberation. Denying moral realism is what separates epistemic objectivists from metaphysical or ontological objectivists.

---

[17] Miller, A. p. 5.
[18] Rachels, J. p. 9.
[19] Miller, A. p. 5 – 6.
[20] *Ibid.* p. 6.

Ontological objectivists accept that objective moral facts or properties exist in the world and are metaphysically real. Moral properties are most commonly thought to supervene on or be reducible to natural properties.[21] There are a variety of different views describing how moral properties fit into the natural world; they are either naturalist or non-naturalist (I cannot go into detail here). An ontological objectivist holds that:

Ethics... [is] objective in the sense that moral predicates – 'good', 'right', and so on – refer to real properties of things. Moral facts are part of the fabric of the world.[22]

Miller describes this view as strong Cognitivism.[23] This view seems to run straight into the ontological and epistemological objections, while weak Cognitivism, though not unproblematic itself, seems to avoid these problems. I shall not adjudicate between the two versions of Cognitivism, nor shall I merely pick one to defend against motivational scepticism. My aim is to argue that a correct account of human motivation and Internalism can be compatible with both forms of Cognitivism. This is because when we are looking at the motivational efficacy of moral belief, we are considering a strongly held belief that the subject takes to be justified, by whatever means, and that a belief of this sort gives rise to a desire-like state in rational agents. The issue of what status moral facts have, be it ontological or epistemic, does not bear on how the good moral agent holds her moral belief, as long as she takes it to be objectively true. I turn now to look at Humean motivation before I analyse and argue for Internalism because my argument for Internalism requires the reader to be familiar with certain concepts that I introduce in the next section.

Humean Motivation

To recap, here are the three theses once more:

1. Ethical Cognitivism: moral judgements of the form "It is right that I Φ" express a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.

2. The Internalism Requirement: if someone judges that it is right that she Φs then, *ceteris paribus*, she is motivated to Φ.

3. Humean Motivation: 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.[24]

---

[21] Miller, A. p. 4 – 5.
[22] Rachels, J. p. 10.
[23] Miller, A. p. 4.
[24] Smith, M. p. 12. The relevance of the added stipulation that belief and desire are distinct existences will become apparent in what follows in the rest of this thesis.

The apparent inconsistency turns crucially on how we interpret 3. Let us look then at how Hume and, following him, Neo-Humeans, interpret Practical Reason. Hume's theory of motivation is a theory of how persons reason practically – that is, towards intentional action.

It is the desire, for Hume, that does all the motivational work and the belief that allows one to know how to satisfy one's desire. What one is motivated to do, then, is a function only of what one already wants to do. But can one not arrive at a desire to Φ merely by having the belief that Φing is right or required? Hume's answer to this is famously, No! In order for the mere belief to result in desire, one of two things must be true. Either there must be (a) a necessary connection between belief and desire, where every time I sincerely hold a belief I must come to hold an appropriate desire as well; or (b) Reason must be able, as Plato claims, to give rise to a new motivation.[25] Hume denies (a) because he claims that there cannot be a necessary connection between distinct existences and that belief and desire are distinct existences. Reason, as a separate faculty of the mind from the passions, must be able to create a new desire independently of any appropriately related desire or want already present in the agent's motivational set (S). One does not have to conceive of Reason as a separate faculty to do this, but can claim, as I will, that we do create desires for ourselves through accepting an obligation, independent of the contents of our S (more on this in chapters 3 and 4). Hume denies (b) as he claims that "Reason is, and ought only to be the slave of the passions, and can never pretend to any office other than to serve and obey them."[26] I consider and reject these arguments in chapter 2.

Williams adopts a strategy similar to Hume in arguing against the possibility of (b) Reason must be able to give rise to a new motivation. Williams argues that rational practical deliberation must start from some element in an agent's motivational set (S), for if it did not, we would have nothing to deliberate from.[27] Williams defines a subjective motivational set broadly to include "desires, dispositions of evaluation, patterns of emotional reaction, personal loyalties and any projects ... embodying commitments[28] of the agent."[29] Williams, like Hume, gives an account of the scope of Practical Reason whereby what one has a reason to do is only a function of what is present in one's S. Williams's account of Practical Reason, although broader in scope than Hume's, does not allow for the existence of external reasons[30] for action. Furthermore, if something is to count as a reason for an

---

[25] Plato. *Republic*. Translated by Robin Waterfield .1994. Oxford University Press. p. 327.

[26] Hume, D. p. 415.

[27] Williams, B. "Internal and External Reasons". In *Moral Luck*. 1981. Cambridge University Press. p. 101 – 103.

[28] The inclusion of commitments in the S will be at issue later

[29] Williams, B. 1981. p. 105.

[30] Williams prefers to talk about external reason-statements (which he thinks are always false) rather than external reasons because he believes that external reasons do not exist. Because I shall argue that external reasons exist, I will retain talk of external reasons.

action, Williams urges, it must figure in a correct explanation of that action.[31] An external reason – one that is a reason for an agent independently of the contents of her S – can never be one that motivates an agent in Williams's account of rational practical deliberation. An external reason cannot then figure in a correct explanation of an action, and cannot therefore be a reason for action at all. Only internal reasons really exist, where an internal reason is one arrived at by deliberation necessarily involving some element in the agent's S.

It is clear that Williams's argument turns crucially on his account of Practical Reason and the role of practical rational deliberation. Practical rational deliberation for Williams consists principally of five activities:

1. Ascertaining causal means to the ends one is motivated to attain
2. Finding constitutive realizations of those ends
3. Harmonizing ends by working out how to combine them
4. Ranking ends, where harmonization proves impossible
5. Fully imagining the realization of ends.[32]

Practical deliberation, then, deals with thinking about and working with (roughly construed) one's desires or ends. But to what extent, if at all, does deliberation *affect* one's desires? It seems deliberation is concerned largely with working out what one's desires *are*, and how strong or important they are relative to other competing ends, but not what they *ought* to be. This difference is, I think, crucial. According to 5, deliberation, through the imagination, can add new desires or take away old desires. Williams argues that an agent "may think he has a reason to promote a development because he has not exercised his imagination enough about what it would be like if it came about" and further that, "positively, the imagination can create new possibilities and new desires".[33] How should we interpret this to be consistent with Williams's overall picture? Suppose that Alice, an actress, wants to perform a certain part in a play. Suppose, further, that this role requires that she change her physical appearance – she needs to put on 20kg. Alice imagines what it would be like to fulfil her end and must imagine what it is like to be 20kg heavier. One of Alice's other ends is to be fit and trim. Through the use of the imagination, Alice will realize that she has to give up or modify one of her ends. She will give up either her desire to play that specific part or to be slim all the time. If Alice prefers to play the role, she will not only give up her desire to be slim all the time, but also she will develop a desire to put on weight. It

[31] Williams, B. 1981. p. 108.
[32] Cullity, G. "Practical Theory". In *Ethics and Practical Reason*. Cullity, G. and Gaut, B. [Eds]. 1997. Clarendon Press. p. 107
[33] Williams, B. 1981. p. 104 – 105.

is important to note that the imagination as an operation of reason only affects her desires relative to other ends that she might have.

Williams claims that the only reasons for action that there are must depend on some element in the agent's S. He believes that practical reasons are hypothetical, not categorical. A reason is categorical if it provides a reason for all persons to act as prescribed independently of the contents of what desires persons happen to have. Practical reasons are hypothetical, by contrast, if they are only reasons for persons to act *because* of the contents of their S.

Smith, for example, attempts to respond to the Humean challenge by giving an account of (b) Reason must be able to give rise to a new motivation. He believes that we can rationally set our desires independently of what desires we happen to have. A rational desire in a certain circumstance, for Smith, is a desire that the fully rational agent would have in that situation.[34]

The Humean theory of motivation is a theory about motivating reasons as opposed to normative reasons.[35] To say that someone has a normative reason to Φ is to say that there is some normative requirement that she Φs, that her Φing is justified from the perspective of the normative system that generates that requirement. Normative reasons should be thought of as truths: A's Φing is desirable or required.[36] A motivating reason is not so much concerned with the desirability or rational justification of Φing, but more with a psychological explanation of why the agent Φs. Motivating reasons are psychologically real and cite the states had by the agent that are explanatory (or potentially explanatory) of why she acted as she did or could motivate an agent to act. It is always possible not to act on such a reason because of some unforeseen event that occurs between the time of the agent recognising the reason and the time to act. Motivating reasons typically state a belief/desire pair (for a Humean they necessarily do so): I ate a steak because I desired to taste steak and believed that by eating this meat, I would taste steak. The big question in this debate is whether believing that I have a normative reason to Φ can give me a motivating reason to Φ where a motivating reason is not *already* present.

We have discussed three different sets of reasons:

1. (a) Internal reason: a reason is *internal* just if it can be reached by rational practical deliberation, which starts from the agent's antecedent motivational set.

   (b) External reason: a reason for an agent that is independent of the contents of her S.

---

[34] Smith, M. p. 151 – 171. Smith's account turns on the controversial claim that the desires of fully rational agents would converge. This claim is, however, very difficult to prove and I therefore distance myself from Smith on this point.
[35] *Ibid.* p. 92 – 95.
[36] *Ibid.* p. 95.

2. (a) Hypothetical reason: reasons for action for each person are thus only *because* of the contents of each person's S. I have a reason to Φ only if Φing serves or furthers some desire that I have.

(b) Categorical reason: a reason is categorical if it provides a reason for all persons to act as prescribed independently of the desires persons happen to have.

3. (a) Motivating reason: Motivating reasons are psychologically real and cite the states had by the agent that are explanatory (or potentially explanatory) of why she acted as she did.

(b) Normative reason: a normative reason to Φ is a normative requirement for one to Φ, that one's Φing is justified from the perspective of the normative system that generates that requirement.

My opponents, that is, sceptics about the scope of Practical Reason, claim that only reasons 1(a), 2(a), 3(a), and 3(b)[37] are reasons capable of motivating persons to act. My burden is to claim that all the types of reasons mentioned above do motivate rational persons. My view of Practical Reason must, therefore, be broader in scope than those of my opponents.

Internalism

We tend to think that morality is practical in that judging that something is morally required motivates one to act according to the content of that judgement and that the motivation is internal to the judgement. A moral judgement[38] motivates us intrinsically, or so it seems, without requiring any extra source of motivation other than the moral judgement itself. If Alex sincerely expresses that he judges that he morally ought to refrain from meat eating, we expect him to be motivated to do so (although this motivation need not be overriding). We would be surprised if he asks us to give him a reason to do what he is (or ought to be) motivated to do (or for that which he has a reason to do).[39] This Internalist intuition, like so many in moral philosophy, is not shared by all philosophers. There are two main sources of disagreement about Internalism. The first is about how to formulate the Internalism requirement: whether moral judgements themselves *motivate* us intrinsically or if moral judgements are supposed directly to provide us with *reasons* for action.[40] There is also disagreement about whether Internalism posits a necessary connection between the holding of a moral judgement and the corresponding motivation or reason for action that judgement gives rise to, or if that connection can be defeated in cases of irrationality. The second main source of disagreement is whether or not Internalism

---

[37] My opponents define normative reasons differently to how I do. I discuss this in the Internalism section and in chapter 2.

[38] Philosophers do not all agree that the correct source of motivation in Internalism is a judgement (they reject Judgement Internalism). Since I am calling a judgement that which someone holds when they take an action to be (morally) right or wrong or desirable or undesirable, the obvious source of motivation that we should consider (in this debate) is a judgement.

[39] Nagel views this instance as unacceptable as Alex would be asking for a reason to do that which he has a reason to do. 1970. p. 9.

[40] I leave these notions deliberately vague for now as I mean only to introduce the broad types of disagreement about the ways in which Internalism should be defined. I will give more precise definitions of the various versions of Internalism later in this section.

about moral judgements is actually true. Those who reject Internalism, Externalists, hold that moral judgements do not motivate directly, but require the presence of an additional source of motivation to lead one to be motivated or have a reason to act according to that judgement – they deny that motivation is internal to moral judgements.

In this section I will, firstly, explain the version of Internalism I intend to defend. Audi names this version, "Rational-Agent-Motivational-Internalism".[41] Secondly, I will show why Rational–Agent-Motivational-Internalism (RI for short) is preferable to other versions of Internalism. Once RI has been established as the most plausible version of Internalism, I will argue that we should accept Internalism (RI in particular) over Externalism.

> Rational-agent-motivational-Internalism (RI): *If an agent judges that she ought to perform (or refrain from performing) an action, she will be motivated to perform (or refrain from performing) that action because she holds that judgement, and, if she is not so motivated, she is practically irrational.*

This formulation is similar to Korsgaard's "Internalism Requirement"[42] and Smith's "Practicality Requirement" or "practicality of moral judgement" requirement[43]. There are some important points to note about RI. Firstly, RI is formulated hypothetically (*If* an agent judges...), which implies that the agent in question must actually hold that judgement[44] – the agent must believe that she is morally required to perform (or refrain from performing) the action in question. This is important because if the agent does not actually hold the judgement, the Internalism requirement is not applicable. The second point to note is what RI does not claim – it does not claim that the motivation to act on the judgement will be overriding. It is perfectly possible for an agent to hold a judgement yet not act on that judgement because of a stronger motivation. This feature makes RI a version of, in Brink's terms, "Weak Internalism"[45].

The third point to note about RI is that it does *not* claim that moral judgements motivate *necessarily*. It is on this point in particular that Smith and Korsgaard agree with RI. This weaker connection between judgement and motivation allows for practical irrationality: depression is the

---

[41] Audi, R. "Moral Judgement and Reason for Action" in *Ethics and Practical Reason.* Cullity, G. and Gaut, B. [Eds]. 1997. Clarendon Press. p. 137.

[42] Korsgaard, C. "Skepticism about Practical Reason" in *The Journal of Philosophy.* Volume LXXXIII, *No. 1* 1986. p. 11 And "The Normativity of Instrumental Reason". In *Ethics and Practical Reason.* Cullity, G. and Gaut, B. [Eds]. 1997. Clarendon Press. p. 237.

[43] Smith, M. p. 60 – 91 and p. 7.

[44] I switch between the terms "moral judgement" and "judgement" because I believe that Internalism claims that a moral judgement is a species of practical judgement and that RI applies not only to moral judgements but to all practical judgements.

[45] Brink, D. *Moral Realism and the Foundations of Ethics.* 1989. Cambridge University Press. p. 41.

example most often cited that breaks the connection between judgement and motivation.[46] The connection is still, however, internal – "she will be motivated … *because* she holds that judgement". The "because" here indicates that what explains the motivation is the judgement itself. Nothing apart from the judgement itself (a full moral judgement involves believing that one has an internal commitment – I elaborate on this point in chapter 4) is needed to provide the motivation and the motivation is provided to, say, refrain from meat eating because the agent understands and agrees that meat eating is wrong.

With these important points in mind, I turn to examine rival versions of Internalism. RI is a version of motivational Internalism rather than reasons Internalism.[47] The difference, simply stated, is that Motivational Internalism claims that moral judgements motivate compliant action whereas Reasons Internalism claims that moral judgements give rise to reasons for action. Reasons Internalism is about normative reasons, not motivating reasons.[48] The claim of Reasons Internalism is, therefore, that holding a moral judgement itself gives one a normative reason to act on that judgement (usually necessarily, but I don't think that a Reasons Internalist must be committed to the this). Reasons Internalism is usually invoked by Instrumentalists (those who hold a Humean theory of Practical Reason) who also usually hold that the connection between moral judgement and motivation is a necessary one. If normative reason were defined as I have defined it earlier in the chapter: a normative reason to Φ is a normative requirement for one to Φ, that one's Φing is justified from the perspective of the normative system that generates that requirement, and normative reasons are often external reasons (though not always), Reasons Internalism would be false. It often happens that someone makes an incorrect or undesirable moral judgement (that homosexuals should be abused, for example). If this judgement is wrong or undesirable (which it surely is), then reasons Internalism, employing this sense of normative reason, is false because this judgement does not give the judger a normative reason to abuse homosexuals.

What sense of normative reason do Reasons Internalists have in mind? A normative reason could be, for the Instrumentalist[49], one that best satisfies one's desires or preferences or, for Williams, a reason that is reached by a sound deliberative route starting from some element in one's subjective motivational set. Williams, however, wants to deny the distinction between motivating and normative reasons (I deal with this in chapter 3) and therefore it is best not to consider him as a Reasons

---

[46] Korsgaard, C. 1997. p. 237 and Smith, M. p. 120 – 121.
[47] Brink, D. 1989. p. 38 -39 and Audi, R. 1997. p. 127 – 131.
[48] *Ibid.*
[49] Hume is of no help here as he denies that Practical Reason is normative in the sense required here because he is a Nihilist about Practical Reason. I say more about this in chapter 2.

Internalist in this sense. The Instrumentalist option is of no help here. The Instrumentalist's Reasons Internalism is false because any interpretation of "normative reason" that is appropriately dissimilar to my definition thereof and that could allow Reasons Instrumentalism to get off the ground is subject to counter-examples. Normative reason, for the Instrumentalist, usually means a reason is justified if it *best* satisfies the ends (a broad interpretation of ends including desires, values and long-term goals is the most charitable interpretation), where this refers to some norm of rationality. It is not true that acting on every moral judgement one makes will always best satisfy one's ends. We can imagine a person whose ends are mostly egoistic but holds the occasional moral judgement, and at times acting on a certain moral judgement she holds does not best contribute to the satisfaction of her ends. The Platonic or Aristotelian[50] response that acting morally is conducive to the good or flourishing life is not open to the Instrumentalist, because this norm of rationality is relative to the ends the agent *actually* has. While the Instrumentalist need not be committed to a desire or preference maximizing view of rationality – or something relevantly similar basing norms of Practical Reasoning on what ends an agent has (I argue this in chapter 2) – other norms of Practical Reason (if there are any) go beyond the agent's actual ends to specify what a normative reason is. I have shown that definitions of normative reasons of this sort (such as mine) are of no use to the Reasons Internalist.

If a non-Instrumentalist adopted Reasons Internalism (although I don't see why she would want to) she would face the same problem. Either she claims that a normative reason depends on what desires an agent has – then her account would be susceptible to counter-examples – or a normative reason is not based on what ends an agent has, then she cannot account for incorrect or undesirable moral judgements, which would obviously not give rise to a normative reason thus defined. Reasons Internalism therefore fails because the only interpretation of normative reasons that could support it – claiming a normative reason depends on what ends an agent has – turns out to render Reasons Internalism false.

The next issue to consider is whether holding a moral judgement necessarily motivates compliance with that judgement or whether we should prefer RI, which claims that this connection can be broken by practical irrationality. If I can show that any practical judgement – that is, a judgement about what I should do in a given situation – can fail to motivate in cases of practical irrationality, then there is no reason to suppose that moral judgements must be necessarily motivating. This is because any version of Internalism claims, as a minimal condition for its being Internalist, that moral

---

[50] Even if Aristotle were an Instrumentalist, he certainly would not claim that Practical Rationality is only a matter of maximizing the satisfaction of the ends one actually has.

judgements are practical, that is, tend to motivate action. A moral judgement is a species of practical judgement.

The view that there is a necessary connection between moral judgement (or practical judgement in general) and motivation – I call this view the "Necessary Connection Thesis" – is the most common and widely held version of Internalism. Motivational Judgement Internalism is simply assumed to accept the Necessary Connection Thesis.[51] Why does the Necessary Connection Thesis obviously seem to be the right way to formulate Internalism to so many distinguished writers such that they appear to believe they can simply assume it? The only reason I can think of is that any form of Internalism must say that one is motivated to act on a judgement that one makes *because* one holds that judgement; the very nature and content of that judgement gives rise to motivation. This is usually understood as a conceptual connection between judgement and motivation. It is part of our understanding of the concept of a practical judgement that it gives rise to motivation. For something to be a practical judgement, once we understand the concept correctly, it just is by its nature motivating – it is a matter of logical necessity. It is this interpretation, I suggest, that lurks behind the Necessary Connection Thesis.

The Necessary Connection Thesis only follows from the supposition that there is a conceptual connection between judgement and motivation if this is understood from a third person perspective. How persons reason practically, however, should be understood from a first person perspective[52]. It will be useful to explain this point by employing Korsgaard's distinction between logical and rational necessity and her example of how this distinction is relevant in theoretical reasoning (which aims at belief as opposed to action or motivation to act). Korsgaard writes:

…let us say that a rational agent is one who is motivated by what I call the *rational necessity* of doing something, say, of taking means to an end, and who acts accordingly. Such an agent is *guided* by reason, and in particular, guided by what reason presents as necessary. A comparison will help to illustrate this point. If all women are mortal, and I am a woman, then it necessarily follows that I am mortal. That is logical necessity. But if I *believe* that all women are mortal, and I *believe* that I am a woman, then I *ought* to conclude that I am mortal. The necessity embodied in that use of 'ought' is rational necessity. If I am guided by reason [insofar as I am rational], then I will conclude that I am mortal. But of course it is not logically necessary that I accept this conclusion, for if it were, it would be impossible for me to fail to accept it. And it is perfectly possible for someone to fail to accept the logical implications of her own beliefs, even when those are pointed out to her. A rational believer is *guided* by reason in the determination of her beliefs. A rational agent would be *guided* by reason in the choice of her actions.[53]

[51] Brink, D. 1989. p. 37 – 75. Shafer-Landau, R. "Moral Judgement and Moral Motivation." 1998. In *The Philosophical Quarterly*. Volume 48, No. 192. p. 353. Audi, R 1997. p. 126 – 129 (he calls this generic-motivational-Internalism).
[52] Korsgaard, C. *The Sources of Normativity*. 1996. Cambridge University Press. p. 15 – 16.
[53] Korsgaard, C. 1997. p. 221 – 222.

The difference between logical and rational necessity is that while it is impossible to violate logical necessity, rational necessity can be violated on pain of irrationality. Logical necessity is a third person kind of necessity and it is perceiver independent: Korsgaard is mortal (according to the syllogism) whether she believes it or not. Rational necessity must be seen from a first person perspective – it prescribes how *I* ought to act (or be motivated to act) or what *I* ought to believe. While it is true that we do say that one ought to believe such and such on pain of irrationality from an observer standpoint, we are assuming that the person has certain other relevant beliefs and is (or should be) aware of the relevant evidence. When we make a normative claim like that above, we are implicitly "putting ourselves in the shoes" of the agent in question. The connection between the relevant states in Theoretical and Practical Reasoning – the rational necessity – must be made from a first person perspective.

The mere having of a desire to eat a chocolate, an appropriate means-end belief that by picking up and unwrapping this bar and putting it in my mouth I will eat chocolate, and a desire to pick up and unwrap the bar and put it in my mouth does not guarantee that I will be motivated to do so. What is required is that *I recognise* that I have these states and that *I recognise* the connections between these states. If this first person recognition were not necessary, all thought and rational activity would be removed from rational decision-making as it would be a purely mechanistic process that is guaranteed just by the having of these states alone. This unacceptable conclusion is avoided by holding that it is necessary (I deny that it is sufficient – the person must be rational) for a person to be motivated that the person recognises the relevant connections between her appropriate states. This recognitional feature implies that Practical Rationality should be understood from a distinctively first person perspective and that the relevant necessity between practical judgement and motivation is rational, not logical, necessity.

This way of understanding Practical Reason leaves it open that the connection between practical judgement and motivation can be defeated in cases of irrationality. One can fail to be motivated to adopt the acknowledged means to one's ends because of depression, apathy or the like.[54] Explaining that the failure of motivation is irrational is a further argument, as Shafer-Landau points out.[55] Although the outlines of this argument have been provided in the preceding section, I will provide an argument explaining exactly why this is irrational in the following chapters. Once we acknowledge that Practical Reasoning is essentially a first person enterprise and the relevant type of necessity between practical judgements and motivation is rational, not logical, we have reason to reject the Necessary

---

[54] Smith, M. p. 120 – 122. And Korsgaard, C. 1997. p. 237.
[55] Schafer-Landau, R.. p. 355.

Connection Thesis and accept RI. With RI established as the best interpretation of Internalism, I can now argue against Externalism and in favour of Internalism.

Most Externalists, in arguing against Internalism, argue against the necessary connection thesis and conclude that because the connection between judgement and motivation is not (logically) necessary, we should accept Externalism. Shafer-Landau, for example, writes that "Motivational Judgement Internalism (hereafter MJI) claims that, *necessarily*, those who sincerely judge actions right are motivated to perform those actions" and that "Motivational judgement Externalism (MJE) is the negation of MJI".[56] Brink claims that Internalism of any form has three distinguishable components and that one of these is that, "moral considerations *necessarily* motivate or provide reasons for action."[57] This form of argument, as should be apparent, does not support Externalism because one can be an Internalist while denying the (logically) necessary connection between judgement and motivation. Brink's specific argument against Internalism is that the Internalist cannot explain the Amoralist. The Amoralist is someone who accepts the existence of moral facts and asks why she should care about such facts.[58] To allow the Amoralist challenge to avoid obvious responses, we should say that the Amoralist does sincerely judge that she ought to return a lost wallet but is, despite this judgement, *completely unmotivated* (not only not motivated to a small degree) to return the wallet.

RI does allow us to conceive of such a person and claim that the Amoralist is practically irrational. The Amoralist is of roughly the same status as one who judges that she should get in her car and drive to the cinema (she desires to watch a movie and believes that by driving to the cinema she will be able watch the movie) and is completely unmotivated to do so. A moral judgement is, for the Internalist, a species of practical judgement. In a similar way to how the would be movie-watcher is practically irrational (she might be depressed or apathetic), the Amoralist is also practically irrational. I do, however, owe an account of the Amoralist's irrationality, but this will be provided in chapter 4 once I have said more about theories of Practical Reason and Practical Rationality and how moral judgements motivate action. The Externalist, similarly, owes us an account of why the Amoralist is not practically irrational. Both of these tasks require an analysis of theories of Practical Reason and Practical Rationality.

Although the standard Externalist argument against Internalism fails because it assumes that Internalism must adopt the necessary connection thesis, a different argument claiming that motivation is external to the holding of a moral judgement could be provided. This sort of argument must claim

---

[56] Shafer-Landau, R. p. 353. Italics, mine.
[57] Brink, D. p. 42.
[58] *Ibid.* p. 46.

that it is a necessary condition for one to be motivated to act on a moral judgement that one has an extra source of motivation (apart from the moral judgement itself). This claim cannot be justified independently of an account of what it is to hold a moral judgement and of how Practical Reason operates, and this is no easy task. In the next chapter I examine theories of Practical Reason and Practical Rationality.

Chapter 2: Practical Reason and Practical Rationality

In this chapter I shall argue that we should understand Practical Reason as broader in scope than merely being restricted to means-end or constitutive reasoning. In particular, I intend to show that Practical Reasoning is, amongst other things, about one's ends. We need an account of Practical Reason that is rich enough to explain the variety of ways that persons think about and plan intentional action. I shall lay out the main types of views that philosophers hold about Practical Reason, ranging from the most restrictive account, found in Hume, to the richest accounts provided by Searle and Brandom. A methodology needs to be set up before I can evaluate these accounts. The question, "How should we think about Practical Reason?" must be answered, therefore, prior to an analysis of the merits of the accounts of Practical Reason that I consider. I turn my attention now to setting up criteria that we should keep in mind and invoke to adjudicate which account of Practical Reason we should prefer.

How should we think about Practical Reason?

Practical Reason, to state a rough definition, is the faculty that involves deliberation in order to decide how one should act. Theoretical reason, by contrast, is the faculty that functions to form beliefs about what we take to be true. Philosophers talk of Practical Reason in a descriptive and a normative sense and often conflate the two uses. We must keep in mind the difference between Practical Reason in the normative and descriptive senses when theorising about what the scope of Practical Reason is.

Practical Reason and Practical Rationality

Practical Reason (descriptive) refers to how persons *do* reason practically, independently of how persons *ought to* reason practically. I refer to the scope of how persons do reason practically (descriptive sense) as what is *possible* for persons to reason practically. Practical Rationality (normative) refers to how persons *ought to* reason (to be rational) when deliberating about how to act. What it is possible for persons to reason with and about, obviously affects how it is rational for persons to reason about action, but the relation is asymmetrical. We cannot decide on a normative account of how we ought to reason practically and infer from this what the scope of Practical Reason is, but we can, and I urge that we must, work out what the scope of Practical Reason is (what is possible) and develop a model of Practical Rationality based upon this.

I shall explain this point further with the aid of an example. Consider the Instrumentalist view of Practical Reason and the Expected Utility (EU) view of Practical Rationality. The Instrumentalist (Humean) view holds that "all Practical Reasoning is means-end reasoning, that is, that figuring out what to do is entirely a matter of determining how to achieve one's goals or satisfy one's desires."[59] This view places a limit on how persons reason practically. The EU account of Practical Rationality claims that rationality is a matter of achieving maximum desire or preference satisfaction. The two views, although importantly related (this model of Practical Rationality usually accompanies this view of Practical Reason) are about two different things: how it is *possible* for persons to reason practically and how it is *rational* for persons to reason practically.

These two views, furthermore, can come apart in two ways. One can hold an Instrumentalist view of Practical Reason and reject the Expected Utility view. One could, for example, hold a satisficing[60] view of Practical Rationality.[61] One could also reject Instrumentalism and still hold EU. One could hold that Practical Reason is not exclusively means-end but also involves constitutive reasoning about, say, what would count as entertainment if one wants to be entertained[62] (many Instrumentalists accept this) and that a central feature of Practical Reason is working out what one's ends actually are[63] and still accept EU. Nihilists, such as Hume, would deny the distinction between Practical Reason and Practical Rationality because they claim that there is no such thing as Practical Rationality. I argue against Nihilism later in this chapter and thereby put this worry to rest. This, then, is the first methodological point to bear in mind on our way to finding an account of Practical Reason and of Practical Rationality.

The functionality of Practical Reason:

In trying to find a descriptive account of Practical Reason we are interested in understanding what it is possible to reason practically about, that is, how Practical Reason functions. In particular, we are interested in the restrictions (I *do not* mean normative constraints) on the operation of Practical Reason. Is Practical Reason merely instrumental? Does Practical Reason involve deliberating about ends and, if so, can reason set its own ends? No one would deny that Practical Reason does operate to determine the means to achieving one's ends, but we want to know if that is all it can do. An adequate account must be able to give a plausible explanation of the role of plans in our practical deliberations

---

[59] Millgram, E. "Introduction". In *Varieties of Practical Reasoning*. 2001. MIT Press. p. 4.
[60] Satisficing holds that rationality is not a matter of *maximising* one's desire or preference satisfaction, but of satisfying them in a way that is *good enough*.
[61] Slote, M. "Moderation and Satisficing." 2001. In *Varieties of Practical Reasoning*.
[62] Williams, B. 1981. p. 104.
[63] Schmidtz, D. "Choosing Ends." 2001. In *Varieties of Practical Reasoning*.

and decision-making as this is central to our experience of how we deliberate about how to act. An explanatory requirement of an account of Practical Reason is that it must explain how plans function in our deliberations leading to action.

The sense of "plan" I employ, following Bratman, is that of "having a plan", being in a certain "state of mind".[64] To have a plan is to be in a state of mind whereby one has decided on an action/course of action to carry out in a certain circumstance. We choose plans to coordinate our activities in the future intra and inter-personally,[65] that is, plans allow me to coordinate my activities over time and coordinate my activities with others. It will be useful here to give an example of a plan.

Suppose that there is a lot that I need to get done this afternoon. I must pay rent, pay lights and water, return a video I took out before it is overdue, buy medicine from the pharmacy, buy dog food and return a book to a friend. Buying dog food, paying rent and lights and water are communal tasks that my housemate, Brendan, and I need to complete. It turns out that Brendan also has certain tasks that he needs to perform this afternoon. He must pick up photos from Kodak, photocopy his lease agreement and buy toiletries. Neither of us owns a car and we must walk around town to complete our tasks. We decide to share the tasks between us according to where each task needs to be performed. I plan to pay rent (communal task), return the book to my friend (my task), buy medicine (my task), pick up photos (Brendan's task) and photocopy the lease (Brendan's task). Brendan plans to pay lights and water (communal task), buy dog food (communal task), buy toiletries (his task), and return the video (my task). We each plan to complete the list of tasks we have agreed upon and meet at the pub at 17:00.

We both have plans about what to do this afternoon. In order to carry out my activities most efficiently, I make another plan, a plan about which tasks to perform in what order and where I should go to do this. I have reasonable expectations of how Brendan will act this afternoon, which are grounded on my knowledge of his plan. How can this be? It must be that Brendan's plan plays some sort of motivational role; plans are *pro-attitudes*.[66] Bratman argues that this role is more than merely a *conduct-influencing* role, but a *conduct-controlling role*, since in the normal course of events I simply execute my plan, that is, follow the course of action I have set for myself.[67] Plans have a stronger motivational role than desires in that they do not merely influence my behaviour, they tend to control it.

---

[64] Bratman, M. "Taking Plans Seriously." 2001. In *Varieties of Practical Reasoning.* p. 203.
[65] *Ibid.* 204.
[66] *Ibid.* 204 – 205.
[67] *Ibid.* 205.

Furthermore, my expectations of Brendan's behaviour would only be reasonable, Bratman urges, if plans have a certain stability.[68]

What is it for a plan to be stable, in the sense that is relevant here? First, in the normal course of events it will not vary in response to ordinary, nonrational bodily processes, unlike my desire to eat and drink. Second, and most important, having a plan will involve a strong disposition *not to reconsider* it except in the face of a problem...[69]

One can fill in the details of the plan or make adjustments in one's sub-plans without reconsidering the plan. One's plans should be consistent with one's other projects and therefore tend to be incomplete such that I can fill in certain details of how to fulfil a plan in such a way as not to conflict with my other goals.[70] This is another crucial difference between plans and other motivators such as desires. It is important to notice that plans motivate action and typically do not have any distinctive phenomenological feel. Bratman points out certain other features of plans according to which we can judge whether a plan is rational or reasonable. I shall not mention these here as my primary focus, in terms of taking plans seriously, is on a theory of Practical Reason that must be set up before we try to take plans seriously in a theory of Practical Rationality.

The Villains of Philosophy:

There are certain individuals that philosophers have conjured up to make their opponents stop and re-think claims about how we make inferences, are motivated to act morally and whether or not a motive to justice is binding on persons. These villains are typically called upon to baffle philosophers in situations where all the steps are in place for a person to come to believe something or to be motivated, but the villains resist. Such characters are not uncommon, but I want to focus on two that keep raising their villainous heads. They are Lewis Carroll's Tortoise and the Amoralist as described by Williams and Brink[71]. I shall go through each of them individually.

Railton explains the tortoise argument:

Achilles entertains an argument:

1. If p then q
2. p
3. So: q.

Carroll's Tortoise asks Achilles whether there isn't a gap in this argument, a missing premise. Couldn't one grant both premises but fail to be driven to the conclusion unless one also granted:

---

[68] Bratman, M. p. 206.
[69] *Ibid.*
[70] *Ibid.* p. 208.
[71] Brink, D. p. 45 – 50. Williams, B. *Morality*. 1972. Penguin Books. p. 18.

4. If [(if p then q) &q] then q

To effect the connection between 1 and 2 and 3?

This seems reasonable to Achilles, on whom it only slowly dawns that he has just reached a regress. For suppose our premises are enlarged to be 1 – 4. Tortoise will cheerfully argue we would need a new premise to effect *their* relevance to the conclusion, namely:

5. If { ([(if p then q) & p] then q) & (if p then q) & p } then q

Were 5 added, Tortoise would notice the need for yet another premise to link 1 – 5 with the conclusion.[72]

We learn that rules of inference cannot be added as premises on "pain of regress"[73], but there is something more interesting going on. Could we imagine a person like Tortoise who cannot see that the conclusion follows in this case? Is this person possible and, if so, what is wrong with him? We could change the argument into a practical syllogism such that the conclusion is an action or judgement about how to act. The tortoise desires to eat a chocolate and believes that by unwrapping a Bar-one and biting it he will eat a chocolate, and refuses to accept that he has any reason to unwrap the chocolate. A theory of Practical Reason must say if the tortoise is possible and if so, must have the resources to ground a theory of Practical Rationality to explain his failing. I discuss the Tortoise on p. 39 - 40 of this chapter.

The Amoralist is someone who, like Alex of chapter 1, professes to accept that a certain act or course of action is morally wrong but is in no way motivated to refrain from performing the action when the opportunity arises. The Amoralist traditionally poses a problem for the external-reasons theorist. One constraint on the concept of an Amoralist, as I mentioned in the previous chapter, is that he must actually believe what he professes and must not be motivated in any way (so not motivated less relative to a stronger desire) to act on that moral judgement. My discussion of the Amoralist will not be found in this chapter, but will be given in chapter 4 once I have further developed my argument for the main claim of this thesis.

Theories of Practical Reason

In this section I aim to map out the main positions philosophers adopt regarding the scope of Practical Reason. These views can be divided into two broad categories: end-oriented accounts and act-oriented accounts.[74] End-oriented accounts seek to show how Practical Reason can select action or courses of action that will contribute to certain ends or results.[75] Act-oriented accounts hold that action

---

[72] Railton, P. "On the Hypothetical and Non-Hypothetical in Reasoning". In *Ethics and Practical Reason.* Cullity, G. and Gaut, B. [Eds]. 1997. Clarendon Press. p. 76. I have changed the numbering of the premises that Railton uses.
[73] *Ibid.*
[74] *Routledge Encyclopedia of Philosophy.* Volume 7. Craig, E [Ed]. 1998. Routledge Press. p. 614.
[75] *Ibid.*

or courses of action are guided by norms or principles and that the task of Practical Reason is to identify appropriate norms and principles that we use to guide action.[76] End-oriented accounts tend to restrict the scope of Practical Reason more than act-oriented accounts do, although an account's being end-oriented does not itself prevent Practical Reason from, as one of its functions, selecting one's ends.

One can, of course, be a Nihilist about Practical Reason. This is a position that has been attributed to Hume.[77] "Nihilism about Practical Reason is the view that there are no legitimate forms of practical inference, and that consequently there is no such thing as Practical Reasoning: appearances notwithstanding, there is no mental activity that counts as figuring out what to do."[78] For the Nihilist, what look like practical judgements are merely expressions of emotion or desire understood as "raw feels".[79] This view is not a view of Practical Reason *per se* but it is sceptical that there is such a thing as Practical Rationality, and claims Practical Reason to be more like a robotic, unreflective process than an operation of reason. Nihilism, therefore, cannot properly be classified as end-oriented or act-oriented.

End-oriented reasoning:

Instrumentalism: Means-end only

The Instrumentalist view holds that "all Practical Reasoning is means-end reasoning, that is, that figuring out what to do is entirely a matter of determining how to achieve one's goals or satisfy one's desires."[80] Practical Reason functions only to determine the means to one's ends and it is not the task of reason to work out what ends one has or ought to have (many 'Instrumentalists' do, however, broaden their account to allow that Practical Reason figures out what ends one actually has – but this move is not, I believe, properly Instrumentalist). Instrumentalism about Practical Reason can be summed up by Hume's famous declaration that "Reason is, and ought only to be the slave of the passions".[81]

---

[76] *Routledge Encyclopedia of Philosophy.* Volume 7 p. 614.
[77] Korsgaard, C. 1986. p. 6 – 15. And 1997. p. 220 – 234. Millgram, E. p. 3.
[78] Millgram, E. p. 3.
[79] *Ibid.*
[80] *Ibid.* p. 4.
[81] Hume, D. p. 415.

Figure 1: Theories of Practical Reason

```
Nihilism                                                    ┌─────────────────────┐
                    ┌──────────────────────┐                │ Act-oriented reasoning│
                    │ End-oriented reasoning│               └─────────────────────┘
                    └──────────────────────┘
                   /              \          ─ ─          ┌──────────┐
         ┌────────────────┐    ┌──────────────────┐       │Coherence │
         │Instrumentalism │    │Non-instrumentalism│      │driven    │
         └────────────────┘    └──────────────────┘       │reasoning │
          /       |      \                                └──────────┘
                  |       \      ┌──────────────┐                ┌────────┐
                  |        \     │Reason can set│                │Getting │
                  |         \    │ends          │                │the world│
         ┌──────────────┐ ┌──────────────┐  └──────────────┘     │right   │
         │Means-end only│ │Means-end and │                       └────────┘
         └──────────────┘ │Constitutive  │                ┌──────────┐  ┌───────────┐
                          │reasoning     │   ┌──────────────┐│Kantian  │  │Desire     │
                          └──────────────┘   │Specification of││Universalism│ │independent│
                                             │ends          │└──────────┘  │reasons    │
                          ┌──────────────┐   └──────────────┘              └───────────┘
                          │Maieutic ends │              ┌────────────┐
                          └──────────────┘              │Inferential │
                                                        │expressivism│
                                                        └────────────┘
```
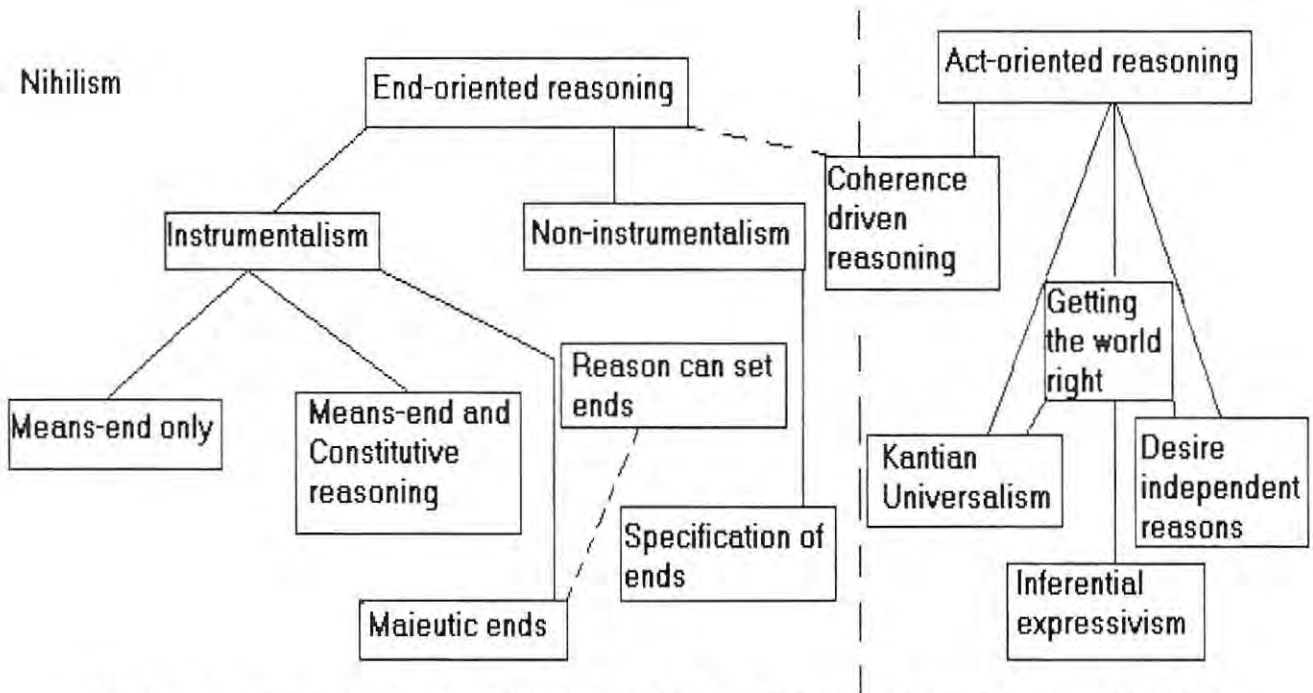
Figure 1 shows the views of Practical Reason that I consider and how they relate to each other.

There are two types of ends involved in instrumental reasoning: ends in themselves and ends that are also means to a further end. Examples of ends in themselves are pleasure, happiness and enjoyment. Being wealthy is an end of human action but is also desired because of the goods that wealth can bring about. One's ends (in themselves) are furthermore "given" as facts about one's motivational make-up and are not rationally decided upon. If one's ends are rationally decided upon it is *only* because they are needed to satisfy some other end that is not itself rationally set.

This account is the narrowest on the spectrum of views as the only form of Practical Reason it allows is means-end reasoning. It does not allow that Practical Reason can evaluate one's ends in terms of justification and consistency (as norms of reasoning); but instead it reasons only about the means to one's ends (or ends that are means as well) in terms of their efficacy to achieving the end. Instrumentalists can allow that reason can assess whether a proposed means to a given end is compatible with the achievement of other ends or means to other ends. Hume can be seen as an example of a (narrow) Instrumentalist, although Millgram and Korsgaard claim Hume to be a nihilist.[82] The Instrumentalist position is one that has been attributed to Aristotle.[83] The next position in the spectrum of views has also been attributed to Aristotle: Practical Reason functions not only to determine the means to one's ends but also to find what activity constitutes the satisfaction of one's ends in a given case.[84] This disagreement stems, in part, from the ambiguity in Aristotle's claim that choice and deliberation are not *of the end* but of *what is towards the end (pros to telos)*.[85] What is towards the end can mean either "means to one's ends" or it can refer to something that counts as a partial or total "constituent of the end".[86] Let us move on to look at the other view that has been attributed to Aristotle.

Means-end and Constitutive reasoning

This view is a broader version of Instrumentalism as it allows that one can reason practically to work out, for example, what would count as entertainment if one wants entertainment. Sometimes it is not appropriate to look only for a means to one's end, as the relation between the act that I need to perform and the end for the sake of which the act is performed is not always a means-end relation. Let us assume that my end is to eat a tasty dinner. I should not always look for some activity that will cause (in a non-philosophical sense) me to have a tasty dinner but often for the activity that constitutes eating a tasty dinner for me. It is not that eating a prawn curry causes me to eat a tasty dinner, but rather that eating a prawn curry is, for me, enjoying a tasty dinner.

This view is still Instrumentalist as it is not committed to claiming that deliberation is also of ends but only of how to achieve one's end, be it means-end or constitutive reasoning.

---

[82] Millgram, E. p. 3. And Korsgaard, C. 1986. p. 6 – 15. And 1997. p. 220 – 234.
[83] Wiggins, D. "Deliberation and Practical Reason". 2001. In *Varieties of Practical Reasoning*. p. 280 – 290.
[84] Kolnai, A. "Deliberation is of Ends." 2001. In *Varieties of Practical Reasoning*. p. 260 – 278.
[85] Aristotle in Wiggins, D. p. 282, and Aristotle. *Nichomachean Ethics*. 2000. Crisp, R. [Ed]. Cambridge University Press. p. 112 – 114.
[86] Wiggins, D. p. 282.

Non-Instrumentalism: Specification of Ends

A major task of Practical Reason, in this account, is to specify what one's ends actually are. We have a set of vague and indefinite ends and need to work out in more detail what these ends are and what it would be like to realise them. Assume that my aim is to write a good thesis. Before I can work out what to do in order to write a good thesis, I need to specify my end, that is, I need to work out what would be a good thesis. Before my instrumental reasoning can get off the ground, the Specificationist argues, we must deliberate about the details of what my end actually is. Williams fits into this category, I believe, because of his insistence on the role of the imagination in working out what the realization of our ends would be like (see chapter 1. p. 12 - 13). Imagining is not, for Williams, the same as deliberation, but an important part thereof. This does not mean that Williams believes that reason is a source of desire, but rather that an important part of practical deliberation is the specification of one's ends. An argument for Specificationism can be found in Kolnai's "Deliberation is of Ends".[87]

Practical Reason, in this view, has two roles. It has an instrumental role as explained earlier and it has a non-instrumental identifying role, that is, it must specify what one's (often vague) ends actually are. Deciding between Instrumentalism and Specificationism is simply a matter of working out whether or not this added aspect of deliberation is a true reflection of what we do when we decide how to act.

Reason Can Set Ends

This is the broadest of the end-oriented views of Practical Reason. There are two ways that I can see this approach proceeding: (a) show how, within a broadly Instrumentalist framework, we set our ends or (b) from a non-Instrumentalist framework, show how we set our ends. Schmidtz uses method (a) by introducing what he terms Maieutic Ends (see figure 1).[88] A Maieutic End is an end achieved through the process of having other ends and, just as final ends are the further ends for the sake of which we pursue instrumental and constitutive ends, Maieutic Ends are further ends for the sake of which we choose final ends.[89] These ends are a type of second-order end (desire) for "filling up affective space".[90] Millgram explains with the aid of an example:

For example, you might want to have a career in medicine, for entirely financial reasons; in order to have the career, you have to care about the right things, e.g., healing the sick, so you come to want to have the end of healing the sick. As the

---

[87] Kolnai, A. In *Varities of Practical Reasoning*.
[88] Schmidtz, D. "Choosing Ends." 2001. In *Varieties of Practical Reasoning*. p. 239 – 241.
[89] *Ibid*. p. 239 – 240.
[90] Millgram, E. p. 9.

example suggests, it is possible to reason about the desirability of wanting something, without expanding the repertoire of inference patterns beyond the instrumental.[91]

The Maieuitic End in this example is the end of healing the sick. Whether or not there really are such things as Maieutic Ends and if it is appropriate to say that we *choose* these ends are matters for debate. I will not deal with them here.

Approach (b) can be attributed to Plato, who argues that Reason can give rise to a new motivation. He holds that Reason both identifies the proper end of action – the Good – and works out how to attain that end. This move, however, threatens the distinction between Theoretical and Practical Reason as human knowledge and desire have a common object – the Good.

Act-oriented reasoning

Practical Reason guides action, in contrast to end-oriented approaches, not according to subjective or objective ends, (it is not teleological) but according to the "practical propositions"[92] (norms, rules and principles) that persons follow in deciding how to live and act. The focus then, is not on restricting practical deliberation to issues to do with means-end or constitutive relations but on finding norms and principles that guide practical inference. The challenge for act-oriented approaches is to find suitable norms and principles and to say why these rather than others guide Practical Reasoning. Let us look at some versions of act-oriented reasoning.

Coherence-driven reasoning

Practical Reasoning involves coordinating one's plans, goals and views of relevant situations in the direction of greatest coherence. One must deliberate about how to achieve one's goals in a way that is most consistent with one's other ends and objectives. Coherence-driven reasoning can be seen in an end-oriented or act-oriented light. It is end-oriented if the search for coherence in one's goals is only for the sake of maximizing one's desire (broadly construed) fulfilment.[93] On this model of rationality, one's set of preferences or ends is either coherent or not, depending on whether it achieves maximum utility, which seems to place too high a demand on human rationality.[94] Perhaps another model of rationality congenial to end-oriented reasoning, satisficing[95] for example, could accompany a coherence-driven account of Practical Reason. Saying that humans strive for coherence in their

---

[91] Millgram, E. p. 9.
[92] *Routledge Encyclopedia of Philosophy*. Volume 7. p. 616.
[93] Millgram, E. p. 12.
[94] *Ibid*.
[95] Rationality is not seen as a matter of maximising utility, but achieving standards that are good enough.

reasoning, however, is invoking a norm or principle of rationality that humans use to guide conduct. It is best, therefore, to think of coherence-driven reasoning as act-oriented.

Coherence-driven practical deliberation "amounts to choosing the subset of goals and actions under consideration that best cohere with one another (and with one's emotions or values, if those are part of the coherence problem). Practical Reasoning can be described as *inference to the most coherent plan.*"[96] If this account is an accurate description of how we reason practically, we would expect persons to employ certain additional norms that assist us in making consistent our plans and goals. This approach requires, of course, an account of precisely what coherence amounts to. An account of coherence-driven reasoning is provided by Thagard, who argues against calculation-based and intuition-based theories of decision-making in favour of a coherence-driven theory of good decision-making.[97] Thagard argues for a computational model of coherent decision-making in which he tries to extract the advantages of the calculation and intuition decision models, while rejecting the disadvantages of these approaches, such that decisions are made on the basis of "an assessment of the overall coherence of a set of actions and goals".[98]

Practical Reason involves "getting the world right"

An important part of Practical Reason (or perhaps all there really is to it) is to learn from experience what is important about how we should act and/or give an accurate evaluation of our situation that is free from our unreflective (often emotional) bias. Practical empiricism claims that "it is both possible and necessary to learn what matters or what is important from experience, and that there is no reason to think that goals, priorities and evaluations and other like pieces of an agent's cognitive equipment will be useful guides to action if the world is not allowed to have its say in what they look like".[99] Murdoch believes that in order to know what to do, one must have a correct description of one's situation and that figuring this out is a matter of Practical not Theoretical Reason as commonly supposed.[100] Her rejection of this common supposition is based on the premise that practical deliberation takes time and often does not arrive at action owing to "curve-balls" that one encounters between the time of deliberation and the time to act.

---

[96] Millgram, E. p. 13.
[97] Thagard, P. "How to Make Decisions: Coherence, Emotion, and Practical Inference." 2001. In *Varieties of Practical Reasoning*. p. 355 – 369.
[98] *Ibid.* p. 362.
[99] Millgram, E. p. 16.
[100] Millgram, E. And Murdoch, I. "The Idea of Perfection." 2001. In *Varieties of Practical Reasoning.*

There seems to me to be some truth in the claims that it is important, for the purposes of practical reasoning, to learn from experience and to have an accurate description of one's situation. This, however, cannot be a complete account. If we accept that persons, when deliberating about how to act, learn from experience and try to give an accurate description of their situation, this does not tell us all that we want to know about the operation of Practical Reason. "Getting the World Right", nevertheless, is an important aspect to be considered in a theory of Practical Rationality (normative sense).

Kantian Universalism

The Kantian position on Practical Reason is that we act on reasons and that good reasons are universalizable. The reason for my acting a certain way in a certain situation, if it is to be a good reason, is one that would be a reason for all rational persons in that situation. Kantian ethics is inextricably tied to a Kantian approach to Practical Reason. We should act only on those maxims that we would wish to be a universal law of nature[101], that is, we ought only to act on reasons that are universalizable. This is one formulation of the Categorical Imperative. Kant wants the foundation of his ethics to be non-heteronomous in that morality cannot be derived from persons' contingent non-cognitive states such as emotions or desires, but from Practical Reason alone. Although moral reasons are categorical, and the supreme principle of morality and accordingly Practical Reason is the Categorical Imperative, Kant admits that instrumental reasoning is indispensable and that hypothetical imperatives issue from instrumental reasoning.[102] The Kantian approach is therefore much broader in scope than the theories I have mentioned previously as it allows both the hypothetical and non-hypothetical in Practical Reason. This approach is an intricate interweaving of the normative and descriptive aspects of practical deliberation. Persons reason according to certain norms and principles (the supreme principle being the Categorical Imperative), but since we are "imperfectly rational" as Kant maintains, we often fail to meet the normative requirements of rationality in our practical deliberations. Kantian principles, the Categorical Imperative in particular, have been criticised for being purely empty formal notions.

O'Neill responds to the charge of empty formalism by claiming that the proposed maxim must neither bring about a conceptual contradiction nor a *contradiction in the will*, that is, one's intending a

---

[101] Kant, I. *Groundwork for the Metaphysics of Morals*. 2002. Hill, T. And Zweig, A. [Eds] Oxford University Press. p. 65 – 76.
[102] *Routledge Encyclopedia of Philosophy.* Volume 7. p. 618.

maxim must be universalizable.[103] We need test not only whether a situation in which every person committed, say, the act of theft would lead to a conceptual contradiction, but also test if a situation in which everyone intended not to give to charity would lead to a contradiction in intentions. O'Neill suggests that if we take the idea of universalizing our intentions seriously we can come up with some substantive requirements of rationality. I shall list only a few of them.

1. We should intend not merely all necessary means to one's intended maxim, but also intend some *sufficient* means to one's intended outcome.

2. We should seek to make the necessary means (one has identified) available when they are not.

3. The various specific intentions we adopt in acting on a certain maxim in a given situation should be consistent.

4. The foreseeable results of acting on a specific intention should be consistent with that underlying intention.[104]

I do not intend to take sides in this Kantian debate now, but merely to provide some insight into how a plausible Kantian account of Practical Reason might look.

Desire-independent reasons for action

John Searle argues against what he terms the "Classical Model of Rationality" (Instrumentalism is largely his target) in favour of a model that allows us to create desire-independent reasons for action.[105] He argues that we create desire-independent reasons for action through the free and intentional imposition of "conditions of satisfaction on conditions of satisfaction".[106] When we place ourselves under a commitment, we are creating a reason for ourselves (Searle concentrates on the performance of speech acts as the tool for doing this). A commitment is, according to Searle, the adoption of a course of action or policy (or other intentional content) where the nature of adoption gives a reason for pursuing the course.[107]

Searle's account, then, is overtly anti-Humean and relies on the formal features (functionality) of our mental states as well as the idea of a free, irreducible self that can set reasons for itself, follow those reasons or be subject to practical irrationality when the Prior Intention (PI) is not transformed into an Intention in Action (IA) – unless there arises some new reason not to act on the PI. A PI is an

---

[103] O'Neill, O. "Consistency in Action." 2001. In *Varieties of Practical Reasoning*. p. 315 – 322.
[104] O'Neill, O. p. 312 – 313.
[105] Searle, J. *Rationality in Action*. 2001. MIT Press.
[106] *Ibid.* p. 173.
[107] *Ibid.* p. 175.

intention formed prior to an action. An IA is the intention I have while I am actually performing the action and can be conscious or unconscious. I discuss Searle in more depth in chapter 3.

Inferential Expressivism

Brandom argues that Instrumentalism is mistaken to ascribe desires to agents as a necessary premise in a practical syllogism because practical inferences do not necessarily involve actual desires, but rather are recast in a form that invokes desires to complete the "material" correctness of an inference.[108] Brandom argues positively that a reason statement that does not invoke an actual desire is not incomplete because a practical syllogism should not be seen as monotonic.[109] We do not reason practically, Brandom maintains, according to the formal features of practical arguments, but according to the content of our mental states and that this is how we form intentions to act. Brandom considers the argument:

It is raining

Therefore, I shall open my umbrella.

Brandom does not treat this argument as an enthymeme (an argument with a missing or suppressed premise) as he thinks that the lack of a premise asserting a desire to stay dry, *contra* Davidson, is not necessary to complete the argument. The inference would not go through, not if I lacked the desire Instrumentalists claim as necessary, but only if I had a contrary desire, say to sing and dance in the rain.[110]

Brandom explains:

But the fact that conjoining a premise incompatible with the desire to stay dry would infirm the inference (turn it into a bad one) does not show that the desire was there all along already functioning as an implicit premise. There would be a case for that conclusion only if the reasoning involved were *monotonic* – that is, if the fact that the inference from p to q is a good one meant that the inference from p & r to q must be a good one. (So that the fact that the latter is *not* a good argument settled it that the former isn't either.)

But material inference is not in general monotonic – even on the theoretical side. It can be in special cases, say in mathematics and fundamental physics. But it never is in ordinary reasoning, and almost never in the special sciences.[111]

---

[108] Millgram, E. p. 18.
[109] Brandom, R. "Actions, Norms, and Practical Reasoning". 2001. In *Varieties of Practical Reasoning*. p. 470 – 472.
[110] *Ibid.* p. 471.
[111] *Ibid.* p. 471 – 472.

This approach to Practical Reason is extremely permissive as it allows that desire is not necessary in order to reach an intention to act. In terms of Practical Rationality, Brandom's account is also very permissive as there could be many different kinds of legitimate inference.[112]

Evaluating theories of Practical Reason

A plausible account of Practical Reason must keep in mind the distinction between Practical Reason and Practical Rationality, explain the motivational role of plans and must answer the villains of philosophy. The obvious place to start my evaluation of Practical Reason is with the narrowest of views: Instrumentalism. I leave a discussion of Nihilism until later. Instrumentalism's distinctive claim is that we do not reason about ends, only the means to our ends or what would constitute the achievement of our ends. The simplest way to refute Instrumentalism involves a negative and a positive argument. The negative argument is to show that there is no good reason to maintain that reasoning is not about our ends. The positive argument must show that we do reason about ends and also how we do so.

Why do Instrumentalists hold that Practical Reason is not about ends or that "Reason is, and ought only to be, the slave of the passions"? I can think of two main lines of argument for this conclusion, one from Hume and one from Williams. Hume argues that reason (belief) alone can never give rise to a motivation. Can we not come to have a desire or end because of a belief that we hold? Hume's answer to this is famously, as I explained in chapter 1, that we cannot.

Hume's account of desires as states that are rationally neutral explains why he believes that a desire cannot be rationally induced. The passions, having phenomenological content essentially and being non-rational, are not states that can be rationally induced. A phenomenological conception of desire, where desires necessarily involve a certain "feel", cannot be married with Instrumentalism, as this simple argument will show.

1. An Instrumentalist account of Practical Reason, employing a phenomenological conception of desire, is bound by the phenomenology of motivation.
2. The phenomenology of motivation is inconsistent with the conjunction of Instrumentalism and the phenomenological conception of desire.
3. Therefore, we need to abandon either the Instrumentalist or the Humean (phenomenological) account of desire.

---

[112] Millgram, E. p. 18 – 19.

To have a desire, in this conception, essentially involves that one feels the "prick of the desire" – that is just what it is to have a desire. Every time we are motivated to act, a desire must be present and we must, therefore, feel the prick of the desire. It is not true that every time I am motivated to act, I feel the prick of a desire. One afternoon I folded my washing and placed it in my cupboard. I deliberated, albeit briefly, about how to fold my washing (I am no expert) and where to put my shirts, pants etc... to make my cupboard most conveniently accessible. During the entire process, from deliberation through to planning my actions to performing my actions, I did not once feel the prick of an appropriate desire. Furthermore, when I make a plan about how to act I do not normally feel some desire pulling me in some direction. In order to take plans seriously we have to acknowledge both that feeling a desire in the sense required above is not necessary to have a plan and that plans do play a motivational role in influencing and even controlling our behaviour. The phenomenological conception of desire, therefore, does not support Instrumentalism.

Hume does, however, recognise that there exist "calm passions" that may appear to lack phenomenological content[113]. This does not do anything to justify the conclusion that desire-like states cannot be rationally induced. If anything, it opens the way for an argument to the opposite conclusion.

Perhaps Williams could be of help here. Williams anticipates the response from the external-reasons theorist that holding a belief can give rise to a new motivation because a rational agent is one who is likely to be motivated to act as he believes he is required to.[114] Williams responds:

What is it then that one comes to believe when he comes to believe that there is a reason to Φ, if it is not the proposition, or something that entails the proposition, that if he deliberated rationally, he would be motivated to act appropriately.[115]

Hooker objects that Williams's response is either question begging or impotent.[116] It is question-begging if by "deliberated correctly" Williams is referring to his conception of rational deliberation in which all deliberation must start from some element in the agent's S (some pre-existing end). [117] This response, however, presupposes Instrumentalism (or at least the claim that we cannot rationally set our ends) in order to defend Williams' internal-reasons theory. The response would be impotent if Williams does not presuppose instrumental rationality and he, therefore, leaves open a view of practical deliberation that allows one to be motivated to Φ by sincerely believing that one ought to Φ.[118] Hooker is spot on here as we have not been given a good reason why Practical Reason cannot operate as the

---

[113] Hume, D. *A Treatise of Human Nature.* Book 2. Part 1. Section 1.
[114] Williams, B. p. 85.
[115] *Ibid.*
[116] Hooker, B. "Williams' Argument against External Reasons." In *Varieties of Practical Reasoning.* p. 100.
[117] *Ibid.*
[118] *Ibid.* p. 101.

external-reasons theorist maintains. It still remains to be shown how believing that I ought to Φ can motivate me to Φ.

We are now pushed into a broader view of Practical Reason, one where we do deliberate about our ends. On a broader view, we only deliberate about how to specify our existing ends and do not set our ends anew. What should we make of the claim that a large task of Practical Reason is specifying our ends? We should agree. I aim to make a tasty dinner for my housemates. Before I can work out the means to this end, I must deliberate about what my housemates would count as tasty – not too spicy, creamy perhaps, but not when mixed with onion and so on. This claim seems to me unproblematic as it conforms to our experience of how we deliberate practically and it is necessary in some situations to specify our end properly before we can achieve it. It could be objected that specifying one's end is the task of Theoretical and not Practical Reason. If the specification of one's end is, in certain situations (like my tasty dinner example), necessary for me to fulfil my end, this is an instance of Practical Reason. It is part of the deliberation involved in deciding how to act.

Specificationism holds that we do not rationally set our ends anew. If an argument cannot be found to this conclusion from within Instrumentalism, I do not see why we should accept this restriction from a broader view without a new line of argument.

We now move to consider end-oriented accounts of Practical Reason that believe that we can rationally choose our ends. Working with Maieutic Ends is an attempt to show how we can rationally set ends within an Instrumentalist account of Practical Reason. The challenge to a defender of Maieutic Ends is to show that these ends are compatible with Instrumentalism; but since I have shown that the demand to stick within the confines of Instrumentalism lacks support, this challenge dissipates. We can, I believe, have ends achieved through the process of having other ends. Persons often come to care about matters that would be beneficial to care about because of certain other ends they had when they previously did not care about those matters.

Plato's view of how Reason can select ends is, however, problematic as his view sets one end as the goals of all reasoning, practical and theoretical. This is problematic for two reasons:

1. Plato's account relies on the existence of a metaphysically objective single end to which all human reasoning aims. Specifying what this end is and showing that it exists is an almost insurmountable task that seems to rely on an implausible anti-naturalist metaphysics. This task may not be impossible, but it is too shaky a ground on which to rest a theory of Practical Reason.

2. It is important to distinguish between Theoretical and Practical Reason in terms of their functionality, that is, Theoretical Reason aims to justify propositions that the believer takes to be true, whereas Practical Reason aims to figure out how we should act. Even if we cannot distinguish between Theoretical and Practical Reason ontologically, the functional difference is indispensable if we are to recognise that our reasoning has different objects – beliefs and actions or intentions to act.

To recap, I have argued that the Instrumentalist has given us no good reason to accept that Practical Reason cannot set its own ends through believing a proposition such as I ought to Φ in certain circumstances. That this belief can give rise to an end or motivation to Φ has not been blocked by Instrumentalism. An account of how the belief in question can give rise to an appropriate desire-like state must still be given. There is good reason to believe that we do deliberate about our ends, especially in the way proposed by Specificationism, as it is often necessary to figure out how to act. My analysis has now reached the border between end-oriented and act-oriented accounts of Practical Reason. I shall now consider the difference in more detail.

The main difference, in the descriptive sense, is that end-oriented accounts claim that all practical reasoning is in some sense directed towards satisfying ends that we have whereas act-oriented accounts claim that we reason according to certain norms and principles, and when we violate these norms it is a failure in Practical Rationality. These two accounts, however, are not mutually exclusive – Instrumentalism needs to posit at least one norm of how we ought to and do reason to be intelligible as a theory of Practical Reason. Dreier's discussion of the villainous Tortoise shows this.

Dreier argues that there is one Categorical Imperative – a principle of rationality that there is a reason for all persons to follow regardless of the contents of their S: the means-end rule.[119]

M/E: If you desire to Ψ, and believe that by Φing you will Ψ, then you ought to Φ.[120]

We can imagine Tortoise agreeing that he desires to taste chocolate and believing that by unwrapping this chocolate bar and biting it he will taste chocolate, but refusing to accept that he has any reason to unwrap the chocolate bar. The lesson Tortoise teaches us is that adding a desire to obey M/E will lead to a regress similar to how adding a belief will not get Tortoise to accept *modus ponens*, but will also lead to a regress. No desire will get you to the means/end rule itself. We must then ask what reason do we have to follow the means-end rule? Adding another desire is futile, but all reasons for the Instrumentalist must contain a desire. Dreier explains:

---

[119] Dreier, J. "Humean Doubts about Categorical Imperatives". 2001. In *Varieties of Practical Reasoning*. p. 92 – 97.
[120] *Ibid.* p. 93.

We can give you practical reasons by finding things that you want, and some things you believe, and getting you to draw practical inferences. If you can't draw the practical inference, not even the fundamental, M/E kind, then nothing can count as a reason for you. This is why M/E has a kind of ground-level normative status. I think it counts as a Categorical Imperative, too. Of course the particular reasons that M/E generates are all hypothetical reasons. But M/E itself is not hypothetical. Its demands must be met by you, in so far as you are rational, no matter what desires you happen to have. This is why I said that I think Humeans are mistaken to say that there are no Categorical Imperatives at all.[121]

If one does not accept M/E as a categorical norm of rationality, we are at a loss about how to explain Tortoise. We cannot say he is irrational because his failing (not obeying M/E) cannot be a rational failing because obeying M/E is not a norm of rationality. To rephrase Dreier's argument in my own terms, Instrumentalism is unable to ground a theory of Practical Rationality if it does not say that persons do employ a norm of Practical Rationality, for if we did not, nothing could count as a reason at all. If there is a norm of rationality, it must be possible to adhere to it. A descriptive account of Practical Reason must claim, therefore, that we do reason according to (at least one) norm(s).

Any account of Practical Reason, if it is to avoid the slide into Nihilism, must employ at least one norm of Practical Rationality. But why is Nihilism false? Nihilism claims, remember, that there is no such thing as a reason for action, but our motivations and actions are merely the expression of our desires and emotions understood as "raw feels". The first reason to reject Nihilism is that it cannot explain what is wrong with Carroll's Tortoise when failing to accept the conclusion of a valid practical syllogism. Secondly, the simple argument I ran on p. 36 - 37 also works against Nihilism's conjunction of the phenomenological conception of desire and belief/desire psychology. It is not true that every time I am motivated to act I must experience some "raw feel". Nihilism, therefore, cannot explain how we come to be motivated to act.

Any account of how Practical Reason operates must claim that we do, when rational, follow at least one norm of rationality. The remaining difference between end-oriented and act-oriented views is largely a matter of which normative account of Practical Rationality to accept. Is Practical Rationality a matter of obeying M/E and maximizing expected utility (some preference-based account for example) or satisficing correctly; or is Practical Rationality a matter of following certain norms and principles governing legitimate inference? On closer inspection, the distinction will grow hazier. This is why I separated end-oriented and act-oriented theories with a dotted line on figure 1. Expected Utility and satisficing accounts implicitly invoke a norm of rationality, albeit a consequentialist norm. It is rational in cases where we aim to fulfil some goal to maximize Expected Utility, or to achieve that goal in a way that we consider to be good enough. This norm is categorical. No matter what your desires or

---

[121] Dreier, J. p. 96.

preferences are, all persons should seek to satisfy them.[122] Why restrict rationality to this extent? The obvious answer is "because the scope of Practical Reason is restricted to instrumental reasoning". But I have argued that there is no justification for accepting this limitation. Because Practical Reason can function to specify our ends and must function to "get the world right" in order to bring about meaningful desire satisfaction, I can see other candidates for norms of rationality on the horizon. Specifying our ends is often necessary to satisfying the goal (especially in a maximizing sense). We can claim that it is a norm of Practical Reason to specify one's ends where necessary as this is entailed, in certain instances, by the EU or satisficing norm. The same analysis will apply to the norms of making one's plans coherent and "getting the world right". If making one's means and ends coherent (seen as eliminating inconsistencies) and "getting the world right" are necessary, in certain situations, to maximize the satisfying of one's preferences or to satisfy one's goals in a way that is good enough, we can derive these norms of rationality from the conjunction of M/E and a norm about achieving maximum desire (broadly construed) satisfaction or satisficing.

Let me return to the example of my and Brendan's plans to explain this point. My end is to complete my tasks and meet Brendan at the pub. In order to achieve this, I must specify this end further. I need to know where to go and when (the shops might be closed at certain times) and I need to "get the world right" to do this. I must learn from experience which route is easiest, how to deal with difficult shop assistants and in short, what the nature of my task is. Unless I do this, I will not complete all my tasks and meet Brendan at the pub on time and therefore will not fulfil my goal. In certain situations the fulfilling of one norm requires the fulfilling of other "sub-norms".

Thus far I have taken the sting out of Instrumentalism and given an outline of what a plausible theory of Practical Reason will look like. Chapters three and four fill in this account and show how my account helps to solve the problem that is the issue of this thesis. I need still to argue how holding a belief that I ought to do something can give rise to a motivation to do that thing without the existence of an appropriately related desire. I will show how my account, once developed further, can explain the villains of philosophy as irrational and what that rational failing is. In the next chapter I discuss the role of desires and beliefs in Practical Reason.

---

[122] At this point external-reasons theorists might start to worry because this rule of rationality makes no reference to the content of one's desires so it would be rational to maximally satisfy your desire for murder, for example. I am aware of this concern and will put it to rest in chapter 4 once I have further explained my account of Practical Rationality.

Chapter 3: The Role of Desiring and Believing in Human Motivation

In chapter 2 I looked at theories of Practical Reason with the distinction between the descriptive and normative interpretations explicitly in view. I want to introduce this chapter by showing how that distinction relates to the distinction I drew in the first chapter between types of reasons for action: between motivating reasons and normative reasons. To recap, to say that someone has a normative reason to return a lost wallet, for example, is to say that there is some normative requirement that she returns the wallet, that her returning the wallet is justified from the perspective of the normative system that generates that requirement. Normative reasons should be thought of as truths: A's returning the wallet is desirable or required. A motivating reason is not so much concerned with the desirability or rational justification of returning a lost wallet, but more with a psychological explanation of why the agent returns the wallet or would return the wallet. Motivating reasons are psychologically real and cite the states had by the agent that are explanatory of why she acted as she did or potentially explanatory of why she could have acted in a certain way. It is, in the sense of "motivating reason" I employ, possible not to act on a motivating reason.[123]

Smith claims that the Humean theory of motivation is a theory about motivating reasons as opposed to normative reasons.[124] Motivating reasons apply, I take it, to how we reason practically (descriptive sense) because a motivating reason is just that reason we can cite that gives a correct description (explanation) of why the agent acted as she did. How we reason practically determines the content of a motivating reason. If, as Hume claims, a necessary condition of practical reasoning is that it must involve a desire, then a motivating reason-statement must refer to a desire. In this chapter I will investigate two claims central to a Humean account of Practical Reason and motivating reasons: 1. In order to be motivated to act in a certain way, it is necessary to have some desire related to acting in that way, and 2. We cannot arrive at a desire to act in a certain way merely by having a normative belief that we ought to act in that way.

---

[123] Audi recognizes this point and prefers to talk about motivational reasons rather than motivating reasons because he wants to emphasise that it is possible for one not to act on a motivating or motivational reason that one has recognised. p. 128. See his footnote 2.
[124] Smith, M. p. 92 – 95.

Normative reasons are related to theories of Practical Rationality (normative sense), but the relation is not as simple as that between motivating reasons and theories of Practical Reason (descriptive sense). To have a normative reason to act in a certain way means that it would be rationally justified to act in that way. Theories of Practical Rationality aim to explain what would make an action (or more correctly a reason for action) rationally justified (to act on). There is, however, a constraint on a "reason" counting as a reason for action (motivating reason or normative reason). Williams writes: "If something can be a reason for action, then it could be someone's reason for acting on a particular occasion, and it would then figure in an explanation of that action".[125] This constraint, as I made apparent in the preceding chapter, is implicit in an account of Practical Rationality because how it is rational to reason practically relies on how it is possible to reason practically. One cannot set a norm for Practical Rationality if that norm prescribes a way of reasoning that is beyond the scope of how people do reason practically. In this instance, "ought" most definitely implies "can". This consideration in turn places a constraint on the concept of a normative reason. A normative reason, if it is to count as a reason for action at all, must be able to figure in a correct explanation of an action, that is, a normative reason must be able to motivate persons to act. Put another way, a normative reason must be able to mirror a motivating reason.

Williams wants to resist the distinction between motivating and normative reasons as I have stated it because he believes that the only reasons for action that exist are internal reasons. An internal reason, remember, is a reason that can be reached by rational practical deliberation, which starts from the agent's antecedent subjective motivational set (S). Williams denies that external reasons exist; where an external reason is a reason for an agent independent of the contents of her S. Normative reasons tend to be (as I have defined normative reasons) external reasons. In order to keep the distinction between normative and motivating reasons, my burden is to show that external reasons that are normative reasons can motivate persons to act as those reasons prescribe. I therefore have to show that the Humean claims 1 and 2 (see the first page of this chapter – p. 42) cannot both be true in conjunction.

I divide the remainder of this chapter into two sections. The first analyses the concept of desire and tries to make clear precisely what is meant by the claim that a desire must feature in all practical deliberation – that is, a reason for action must cite an element in one's S. I then turn to look at the role that beliefs play in human motivation. In particular I investigate whether believing that I ought not to eat meat can give rise to a desire not to eat meat when such a desire was not already present.

---

[125] Williams, B. 1981. p. 106.

The role of desiring in motivating action

Suppose Alice finds a lost wallet in a restaurant and judges that it would be morally right for her to return the wallet. There is further, we can assume, a normative reason for Alice to return the wallet. Suppose that Alice did in fact return the wallet because she judged that it was the right thing to do and was thereby motivated. The Humean has to explain Alice's situation by referring to some want or desire she has. Therefore, because Alice did return the wallet, she must have desired or wanted to do so. The Humean must refer to an element in Alice's S related to returning the lost wallet, and claims that this element cannot be arrived at only through having a normative belief that she ought to return the lost wallet.

There are two broadly different ways that "Alice *wanted* to return the wallet" can be interpreted. In the first sense, it is not true that everything we do is done because we want to. It is perfectly intelligible for me to say that I kept my promise to help my housemate with her work, not because I wanted to, but because I had to or believed that I had to. But that does not imply that there is no sense in which I wanted to keep my promise. If I did indeed keep my promise, it is because I did, in a sense, want to. Nagel recognises this point when he recognises that there is a sense of "want" in which, if I intentionally perform an intentional action, then it simply follows that I wanted to perform that action.[126] How should we understand the difference and relationship between these two senses of "wanting to do something"?

In one sense, "want to" refers to some preference, desire, something that one would enjoy doing, benefit from doing, something one wants to do. The explanation, "because I had to" (or believed I had to) refers to an obligation, responsibility or commitment. In the second sense of want (desire) where everything I do must be because I want to, a want is a motivation – I only do that which I am motivated to do. But if it is true both that everything I do is a function of what I want to do (what I am motivated to do) and that the explanation, "I did it not because I wanted to, but because I (believed that I) had to" is an intelligible explanation, it follows that I "had to" means that I "want to" in the second sense.

There is therefore a need to distinguish between two senses of sets of desires or wants. Schueler distinguishes between "pro-attitudes" and desires proper.[127] Although he adopts Davidson's term "pro-attitude"[128], he defines "pro-attitudes" in a slightly different way from Davidson. I will not look at the difference between the two interpretations, but work with Schueler's definition. He writes:

---

[126] Nagel, T. p. 29.
[127] Schueler, G. *Desire: Its Role in Practical Reason and the Explanation of Action.* 1995. MIT Press. Chapter 1.
[128] Davidson, D. p. 3 –19.

I will regard it as analytic that someone who intentionally performs action *a* under description *d* has a pro-attitude toward performing *a* under description *d*. That will allow us to reserve the term "desire" itself for cases where it *is* possible for someone intentionally to perform some action *a* under description *d* *without* having any desire, in this [desire proper] sense, to perform *a* under description *d*.[129]

Schueler denies that when we have a pro-attitude we will also always have desire proper as well because he believes that I can be motivated to act by some pro-attitude that is not a desire proper such as a belief that I have a responsibility to do something or a belief that I have some commitment.[130]

Does the Humean mean pro-attitude or desire proper when she claims that a desire is necessary for motivation? I shall consider each option in turn. Assume that it is a necessary condition that when I come to form an intention to act after a process of rational practical deliberation that I must have some pro-attitude towards performing that action. Let us consider this in terms of motivating reasons first. "Pro-attitudes" is a broad definition that claims that, as a matter of analyticity, every action requires the existence of a pro-attitude. It is not a substantive notion saying what states can count as pro-attitudes. There is no reason in this account of pro-attitudes to suppose that a belief about what I ought to do (say that I have a commitment to help my housemate with her work) cannot be a pro-attitude. Allowing this, however, takes the sting out of the Humean challenge as it denies the claim (chapter 1, p. 5) that no belief can, on its own, motivate one to act. To avoid this possibility, the Humean needs a substantive notion of desire.

We must therefore attribute to the Humean the view that a desire proper is a necessary condition for motivation. How should we define desire proper to allow the Humean claim to be true? The first option is to adopt the phenomenological view of desire – a view often attributed to Hume. Desires cannot be said to be either rational or irrational apart from insofar as they are based on false beliefs. For Hume, desires are non-rational states that are a species of the passions, which are characterised by a certain phenomenological feel. Desires, like sensations, are simply and essentially states with a certain phenomenological content.[131] This conception of desire cannot be of use to the Humean as I showed with my simple argument in chapter 2 (p. 36 - 37).

As I remarked in chapter 2, Hume does recognise the existence of what he calls "calm passions" that are sometimes called Reason (incorrectly Hume would surely maintain).[132] This suggests to me that Humeans need not be committed to holding that all desires proper have phenomenological content essentially. Fehige puts forward an account of desire designed to overcome the problems associated

---

[129] Schueler, G. p. 35.
[130] Schueler, G p. 35.
[131] Smith, M. p. 104 – 105.
[132] Hume in Mackie, J. p. 31 – 32.

with purely phenomenological definitions but that still gives a central role to how desires make the agent feel – he calls this the "Hearty View".[133] Fehige uses a basic notion of desire that he attributes to philosophers such as Aristotle, Descartes, Kant, Hobbes and Mill, where desires are pleasant thoughts – they are affects.[134] All desire must be accompanied by (or be constituted by) some pleasant thought at the fulfilment of that desire. Things close to our hearts (Fehige's desires) are "projects, ends, goals, or purposes, the contents of our pro-attitudes and inclinations: they matter to us; we care about them, we appreciate, cherish, desire, prefer, value, want and, wish them."[135] Fehige intends his account to be of desires proper and not pro-attitudes although it seems to me that he does, at times, slip between the two uses. He claims that his view is not revisionist, and by revisionists Fehige means, "… those authors who saw off the phenomenal part of desire. They leave us with a torso of the concept, with a behavioural persiflage of desire. Their desire is desire as instantiated in robots, or in thermostats."[136] We had better not then, on pain of cynical reprimand, attribute the pro-attitude view to Fehige.

Fehige wants to break the necessary link between desires and phenomenology that is the downfall of the phenomenological conception of desire. He does this by claiming that it is not necessary that we always do feel pleasant thoughts towards fulfilling a desire for that state to count as a desire, but that Mary desires p if and only if *"in a cool hour, when she's sober, awake, undisturbed by other thoughts,* Mary would be happy [be pleased] fully representing to herself that *p*."[137] There are a lot of seemingly *ad hoc* mitigating factors to allow Fehige. For one thing Fehige is committed to saying that persons constantly in a drunken or depressed state do not have desires. These people, although irrational or mentally unwell in some way, surely do have desires. Let us, however, be charitable to Fehige and look past these problems. I am sure a similar account (or response from Fehige) can easily be found that can overcome this objection.

Fehige's account and any other affective account of desire (that adopts a broadly Instrumentalist account of Practical Reason) face a seemingly insurmountable problem. This problem stems from the incompatibility of two claims that Fehige holds. The first is that to desire that p essentially involves, other things being equal, having pleasant thoughts towards that p. The other is the Humean claim that having an appropriately related desire is necessary to motivate intentional action. The conjunction of these two claims is incompatible with a fact about rational human action: I can and do perform some actions grudgingly without having a pleasant thought about the action (or about any motivation

---

[133] Fehige, C. "Instrumentalism." 2001. In *Varieties of Practical Reasoning.*
[134] Fehige, C. p. 51 – 52.
[135] *Ibid.* p. 53 – 54.
[136] *Ibid.* p. 52 – 53.
[137] *Ibid.* p. 51.

necessary for the action) I performed. I recently ate a stick of my friend's Springbok dry wors on the agreement that I would refund her the cost of the dry wors the next day. She told me the next day that the dry wors cost R15, a figure I considered far too high. As it later turned out, she was mistaken and the dry wors cost only R3,50. I deliberated about what to do and decided to pay her the R15. At no stage in my deliberation did I have any pleasant thoughts whatsoever (as the reader, I'm sure, could imagine) towards paying her the R15 or about any motivation that led me to do so - in this case, my belief that I had a commitment to fulfil.

Fehige would be at a loss to explain this situation. None of Fehige's mitigating circumstances applied and yet I performed an intentional action with no pleasant thought at all regarding the action or any motivation to perform that action. It was not the case that I had some indirectly related pleasant thought towards paying back my friend or fulfilling a commitment. The affective conception of desire, even allowing Fehige to escape the problems associated with strictly phenomenological conceptions of desire, cannot be married with the Humean account of motivation.

I want now to look at Williams to see how he interprets desire and maybe this can revive the Humean challenge. Williams talks of desires not only in terms of motivating reasons but also in terms of normative or justifying reasons. He wants, remember, to deny the distinction. Williams counts as necessary for *rational* practical deliberation that deliberation must involve some element of the agent's S. Williams, in an often quoted passage, defines the contents of one's S to include, "dispositions of evaluation, patterns of emotional reaction, personal loyalties and various projects, as they may be abstractly called, embodying commitments of the agent."[138] There is no doubt that Williams defines desires broadly. Does Williams, when talking about the contents of an agent's S (desires) mean pro-attitudes or desires proper? If Williams means pro-attitudes, it is not clear how he can claim that external reasons do not exist as pro-attitudes do not exclude one's belief that one has a commitment to act in a certain way. Schueler writes:

... pro attitudes cannot possibly be the motivating factors that the internal account [Williams's internal-reasons theory] claims are required for something actually to be a reason for someone to do something if we want to maintain there is a genuine difference between an internal and an external account of justifying reasons.[139]

Schueler's argument to this conclusion is based on the claim that Williams's account is one of justifying reasons – ones that give us a good reason to act. But if any action must be motivated by a pro-attitude and if Williams is talking about pro-attitudes, then Williams cannot make the distinction between justified and unjustified reasons for action – all reasons for action are good reasons.[140] This

---

[138] Williams, B. 1979. p. 105.
[139] Schueler, G. p. 51.
[140] *Ibid*.

argument is, however, somewhat unfair to Williams because he wants to deny the difference between normative and motivating reasons as I define them, which is how Schueler uses them. Williams is giving an account of how Practical Reason (descriptive sense) operates – it must involve some element of the agent's S. It seems that on Schueler's interpretation, it is consistent with Williams's thesis to say that we can be motivated in the absence of some element in our S, but only irrationally so. I do not think that Williams would accept this.

Williams is claiming, rather, that the only types of reasons for action that exist are internal reasons (hence his refusing the distinction between motivating and normative reasons). Whether a reason is good or not depends on the particular deliberative route that gives rise to that reason. I see no reason why Williams could not claim that a bad reason for action is one that, for example, conflicts with other ends of the agent and this was not realised in the deliberation, hence the irrationality. In short, Williams is giving an account of Practical Reason and not of Practical Rationality (although he hints at what this might look like). Williams claims that the presence of an element in one's S is *necessary* for rational, practical deliberation and not that it is *sufficient*. Schueler's mistake is not to take seriously enough his own recognition that Williams rejects the distinction between normative and motivating reasons.

Despite this confusion it still remains that Williams is talking about desires proper and not pro-attitudes because, apart from anything else, it is clear that Williams wants a substantive notion of the S. The question now arises: what do all the states that form part of the S have in common that classifies them as desires proper and allows them to play the distinctive role that Williams supposes in Practical Reason? I have argued against both a phenomenological conception of desire and an affective conception even when it is free from the problems burdening the phenomenological conception. The common element that leads to failure for both of these accounts is that they claim that desires are more like sensations than feelings where desires can be both. Following Millgram, feelings and sensations can be distinguished as follows:

...it is a central feature of sensations that having them involves being aware of them. Feelings, however, may be had unawares; a familiar example is the depressed person who does not realise that he is depressed. Unlike sensations, the feeling of pleasure may sometimes be recognized only in retrospect, or when attention is called to it by others. Feelings can *involve* sensations, sometimes in a way that makes it tempting to say that the sensations are part of the feelings; we often identify feelings by the sensations they involve (that sinking feeling). But feelings do not always involve the same sensations: a feeling of elation may be accompanied by a sensation of light-headedness in the one case, by a feeling of

butterflies in the stomach in another, and by no special sensation in the third; and I may be too intent on what I am doing to notice that I am elated.[141]

We do, however, need an account of desires proper that does not disallow that some desires, say a desire to eat, are sensations. The account must be broad enough to include all the states that Williams lists as part of an agent's S, but must not be so broad as to assimilate desires proper with pro-attitudes. The account must also provide a common factor between, to borrow Nagel's terms, motivated and unmotivated desires.[142] The easiest way to understand this distinction is to consider that we have some ends that are ends in themselves (hunger, thirst, lust) and some ends that are desired because of their relation to other ends (be it means-end or constitutive). Unmotivated desires simply assail us without the need for rational deliberation, whereas motivated desires do not. The latter are typically arrived at via deliberation.[143] This explanation is only slightly different in detail from Nagel's, but retains the spirit of the distinction. The account must be able to explain the role (necessary or not) that desires play in motivating intentional action. The account of desires proper must further recognise that desires are states that are psychologically real – they are real states had by agents.

I suggest that the best way to describe the set of states classified as desires proper is in terms of their sharing a direction of fit. Desire has the world-to-mind direction of fit[144] in that the world should accord with how one desires it to be. A desire does not aim to represent how the world is, but rather is a state that aims to change the world to match the content of the desire. The condition of satisfaction of desire is not that it be true (desire is not the type of state susceptible to truth judgements), but that it be fulfilled. Desires are either fulfilled or frustrated depending on whether or not the world changes to match how one desires it to be.

Talk of direction of fit is, as often noted, purely metaphorical. This is not, however, a problem. The metaphor is intended to show what desires proper have in common. These states share a defining formal feature in that they form a proposition with a specific structure: A desires that p where p refers to a state of affairs such as that the government ban the selling of meat products; or A desires to Φ where Φ is some action such as to eat a vegetarian meal. A state had by an agent is a desire proper if and only if the state can be described in the form: A desires that p or A desires to Φ. What these two formulations have in common is that they are directed towards some outcome (an action or state of affairs) that *should* be the case.

---

[141] Millgram, E. p. 332 – 333.
[142] Nagel, T. p. 29.
[143] Schueler, G. p. 21.
[144] *Ibid.*

Fehige objects to the direction of fit account on the grounds that when we interpret "should" either "statistically" or "normatively" we see that this definition of desires is problematic.[145] If "should" is read normatively, Fehige claims, then what makes a state a desire is that for people who have that state, it is rational to change the world to match the content of that state. Schueler then assumes that interpreting desires in this way, "putting desires to the service of rationality, we will get a notion of rationality based on rationality. We will have gone full circle: it is rational to do what it is rational to do."[146] There is no reason to assume, as I showed by distinguishing theories of Practical Reason from theories of Practical Rationality in the previous chapter, that one who defines desires in terms of direction of fit has to hold a view of Practical Rationality based on desire/preference fulfilment. I need not hold that the standard by which we (always and only) judge Practical Rationality is according to whether or not one fulfils one's desires. Moreover, Schueler is reading too much into the word "should". The word "should" points out an important feature of desires: that desires are about actions or states of affairs that the agent takes not currently to be the case. A state is a desire if, and only if, that state is *directed towards* some action or object that the agent believes is not currently the case.

Desiring some state of affairs that *p*, a second objection runs, threatens the distinction between desire and belief in terms of their direction of fit because, similarly to desiring that the world is free of meat products, a belief is about a state of affairs that can be true or not – a belief has the same structure as these propositional desires. Just as I can desire that the government ban the selling of meat products, so can I believe that the government *has banned* the selling of meat products. Notice, however, that on closer inspection the formal structure of the belief and desire mentioned above is relevantly different. The content of the proposition believed is supposed to be true whereas the content of the proposition desired is supposed not to be the case – when we desire we desire a change in the world. This feature of desires is supported by the fact that when we realise that the state of affairs that we desire is actually the case we lose that desire. Typically then, though not necessarily, we tend to be happy about the state of affairs when we realise that it is the case. Being happy that South Africa beat Nigeria at soccer is a state with a zero direction of fit.[147] This is not a problem, as emotions such as the one described above

---

[145] Fehige, C. p. 56.
[146] Fehige, C. p. 56.
[147] Searle, J. p. 38.

do not directly play a role in motivation. They play a role indirectly in giving rise, through deliberation, to a desire of the form I desire to Φ.[148]

In order to play a role in deliberation and motivation, desires (as I have defined them) must be states had by an agent – they must be psychologically real. This added condition is necessary if we take seriously the truism implicit in Williams's account (as we surely must) that in order to reason in the mind on the basis of a reason, an agent has to have a reason in mind to reason from.[149] The direction of fit account of desires is broad enough to include all the different types of desires we have as it defines desires according to their common features. My account is not as broad as the account of pro-attitudes as it tells us more about desires than that they are states that must feature in motivation. It sets a necessary structural feature for a state to be a desire.

The role of believing in motivating action

In this section I shall focus on the second Humean claim that I mentioned on the first page of the chapter: we cannot arrive at a desire to act in a certain way merely by having a normative belief that we ought to act in that way. I shall argue that this claim is false by showing that we do, in fact, arrive at a desire (as I defined it earlier) to act in a certain way by believing that we are required to act in that way. These types of desires that can be arrived at via a belief (without the assistance of another appropriately related desire) that one has a commitment or obligation. I shall argue that a rational agent that holds a moral belief about how she should act in certain situations will be motivated to act as that belief prescribes because such beliefs give rise to a commitment in rational agents – and commitments are desires on the direction of fit account.

First, however, I need to make some general remarks about the nature of beliefs and what is involved in holding a moral belief. Belief is distinguished from desire as it has the opposite direction of fit. Belief has the mind-to-world direction of fit[150] where the world should be as one represents it to one's mind. That is, one's belief aims to be a true representation of how the world is. The condition of satisfaction of belief is that it is true – beliefs are states that are susceptible to judgements of truth. When we consider if belief in a moral proposition is sufficient to motivate, we are looking at a belief of

---

[148] It could be objected that I can, for example, desire my office door to be closed even when I believe it is closed. This objection challenges my claim that one loses a desire when one realises that the object of desire is currently the case. Even if this objection is correct, which I doubt, one must realise that a "desire" for the door to be closed when I believe it is closed has a zero direction of fit – it is a state that is not directed towards a state of affairs that should be the case, but towards a state of affairs that is the case. The "desire" (I would call it a state of being pleased) does not directly play a role in motivating action directly – it could give rise to a desire (with the world-to-mind direction of fit). This objection does not, therefore, damage my account.

[149] Searle, J. p. 214.

[150] Ibid. 38.

the form, "I believe that I ought to (it is right that I) only eat vegetarian food". It is not just the belief that vegetarian eating is right or required, but that this is required of *me*. The belief at issue then is a self-referential belief as it refers to something that *I* should do. Further, it is a belief about what I should *do* as opposed to merely think or believe. The appropriate moral belief then is a belief about what *I* should *do*.

This difference is important. A sceptic about the motivational efficacy of belief typically asks, "how can recognising a fact to be true motivate one to act? I believe that *x* is a desk. So what! Why is there any link between the recognition of that fact and my motivation to act?" The beliefs at issue (self-referential beliefs) are similar to the belief that *x* is a desk in that they are supposed to represent a fact about the world (construed broadly to allow for epistemically objective facts as well as metaphysically objective ones), a fact about what it is right for me to do. Moral beliefs (the ones that I am considering here) are different from the belief that *x* is a desk as moral beliefs are about what it is right for *me* to *do*.

Beliefs, like desires, can be stronger or weaker. I can have different types of beliefs: weak beliefs, strong beliefs, probabilistic beliefs and beliefs that are sincerely held (to list just a few). A strong belief differs from a weak belief in the degree of certainty with which it is held. Strong beliefs are typically more justified than weak beliefs. A probabilistic belief is one where I take it as more likely that the proposition in question is true than false. A weak belief is often a probabilistic belief, but not necessarily. I believe that the ANC will win the next election. I am very confident that the proposition, "the ANC will win the next election" is true, but I am confident because the likelihood is high (probability of the proposition being true) that the ANC will win.

A sincerely held belief is typically a strong belief, but it is also a belief that I take to be of great significance to my life. I am confident that the proposition "*x* is a desk" is true, but I do not take that belief to play a significant role in how I see my role in the world. A sincerely held belief is an evaluative belief or has an evaluative component. Alice's belief that God exists is an example of this. Although it is true that Alice's belief that God exists need not be a strong belief, the strength of the evaluative commitment depends on the strength with which the belief is held. Whether or not a belief has an evaluative component depends on whether the agent takes the content of the proposition believed to have an impact on how she ought to live or not. Moral beliefs are examples of sincerely held beliefs.

Are there such things as sincerely held beliefs that involve an evaluative commitment, that is, put more interestingly, can sincerely believing a moral proposition give one a reason to act in a certain way in a certain situation? Internal-reasons theorists answer "yes", but not because of the nature of the belief itself, but because of the contents of one's S. It seems odd to me that sincerely holding a belief

52

cannot influence the nature of one's motivational constitution if that constitution does not already explain the influence. It is odd for a number of reasons:

1. The desires a person has can affect how that person forms a belief.
2. Strong belief is itself a concept that entails other commitments for the agent.
3. I can hold a belief about what I should desire and it seems odd that I cannot come to desire the object in question if I don't already, directly or indirectly, desire it.
4. The sincere beliefs that a person holds play a large role in how she views the world and acts in the world.

Point 1 is true if the phenomenon of motivated believing is a real phenomenon. Suppose that there is sufficient evidence that Bob has considered that supports the conclusion that his wife is cheating on him. Despite the evidence, Bob refuses to believe in his wife's infidelity. It is not because the evidence is flawed in any way, but because Bob does not want to believe that his wife is a cheat. This strong desire of Bob's, that his wife be a good and faithful wife, distorts the way he interprets the evidence. We say that Bob is irrational (in a theoretical sense) in not changing his belief to meet the evidence and that he should change his belief and we know that he is capable of doing so. It seems odd that in a case where I believe that I ought to desire to Φ, but do not, in fact, desire to Φ, I cannot be called practically irrational (Hume holds this view, contra rationalism) and that I cannot change what I desire.

Searle argues for 2 as he claims that we (*contra* Hume) derive an "is" from an "ought" wherever we turn.[151] Searle argues that the nature of speech acts (when sincerely asserting a proposition) has certain logical entailments that can give rise to commitments in the speaker. I shall not consider his arguments here, but apply a similar analysis to sincerely held self-referential beliefs (Searle does not expand his account to include beliefs as I do). My arguments can be assessed independently of Searle's and therefore do not presuppose the truth of Searle's claim (apart from those claims I argue for). Holding a belief, in itself, logically prescribes certain other beliefs that we ought or ought not to hold, on pain of irrationality. If I believe that x is a desk, then I ought not to believe that x is an elephant. Similarly, if I believe that petrol has an alcohol content higher than 70% and I believe that liquids with an alcohol percentage above 70% cannot be frozen, then I ought to believe that petrol cannot be frozen. Notice that this entailment is independent of the truth of either of the claims, but derives from the perceived truth (they are believed) of the two preceding claims. As a matter of fact, the two beliefs that give rise to the rational entailment: petrol cannot be frozen, are false. The point is

---

[151] Searle, J. p. 182 – 183.

that if I hold those beliefs, I ought to believe that petrol cannot be frozen. The question of the falsity of the two claims and whether or not I have reason to believe them is a separate matter. Holding certain beliefs, then, gives the believer reasons to hold certain other beliefs: reasons that must be followed on pain of irrationality.

Now consider 3: I can hold a belief about what I should desire and it seems odd that I cannot come to desire the object in question if I don't already, directly or indirectly, desire it. This has the odd consequence that no matter how sincerely I believe that I ought to desire something, and don't already desire it, I might as well not have that belief – it does nothing at all. This is very odd given 4, which will be explained presently. Moreover, the claim that I ought to have such and such a desire loses a lot of its sting if I do, even indirectly, already have that desire. My claim that you ought to do *x* is trivial if you are busy doing *x* or have just done *x*.

To hold a belief is, as 4 claims, already to view the world in a certain way. A good test of whether or not a person actually holds a belief that she professes to hold is that, when the appropriate conditions arise, she acts on that belief. According to the Humean, if I believe that I ought to return a lost wallet in appropriate circumstances, I will only be motivated to return a lost wallet in appropriate circumstances if I already have a desire in my S related to returning the lost wallet when appropriate. When I express the belief that I ought to return the lost wallet when appropriate and fail to do so in circumstances I judge to be appropriate, this means one of four things:

(i)   I do not sincerely believe that proposition.

(ii)   I am practically irrational

(iii)   I acted on another desire, one to do something other than return the wallet, owing to my judging that I have a better reason to perform a different action, or owing to weakness of the will or the like.

(iv)   There was no element in my S relating to returning the wallet in appropriate circumstances.

According to the external-reasons theorist, if (i) – (iii) do not hold, then I will return the lost wallet. The internal-reasons theorist denies precisely this. The internal-reasons theorist must deny (iv) as it stands: the sincere beliefs that a person holds play a large role in how she views the world and acts in the world. The internal-reasons theorist must adapt (iv) to read: (iv: a) the sincere beliefs that a person holds play a large role in how she views the world and *acts* in the world, if and only if she has a pre-existing desire in her S appropriately related to the content of the sincerely held belief. On this reading, however, (iv: a) does not seem particularly meaningful and it changes the test for whether someone actually holds a normative belief that she professes into a test of whether or not a person has a certain

element in her S. The test is no longer for the sincerity of the expressed belief, for in the absence of the sincerely held belief the person would still be motivated to return the lost wallet in appropriate circumstances because there is a pre-existing desire in her S regarding returning lost wallets when appropriate. If she has a pre-existing desire in her S regarding returning lost wallets, she is likely to form some means-end belief about how to return the wallet when it appears that appropriate circumstances might be instantiated.

A concept of belief that excludes the existence of external reasons has, I have argued, some odd consequences, some consequences that give us good reason to reject such a notion of belief. This having been said, it still remains for me to show how believing that I ought to return a lost wallet in appropriate circumstances (in the absence of an appropriately related element of my S) can bring me to be motivated to return the wallet. I turn to fulfil this burden presently.

For someone to say of me that I have an obligation or commitment to become vegetarian, when I do not already have any desire related to becoming vegetarian or have not accepted that I have this commitment, is to say that I have an external reason to become vegetarian. If this demand is normatively justified, it is a normative reason as well. Once I accept that I have this commitment to become vegetarian – once I accept it as binding on myself – I come to have a commitment to stop eating meat. Having a commitment to refrain from meat eating is having a desire to refrain from meat eating as a commitment is a desire on my account of desires (I explain this further in chapter 4). As soon as this happens, I have an internal reason to become vegetarian. An external reason plays a role in motivation when it becomes an internal reason and this occurs when the agent accepts the external reason, that is, accepts the commitment as binding on herself. That the (external) reason was a normative reason (if it is normatively justified) does not change when the reason becomes internal, for there is no reason to suppose that a reason loses its normative force when it is accepted by an agent. Let me explain how all this works.

I have used the term commitment in two different senses – one external and the other internal. It is external when someone says of me that I ought to have the commitment and this is justified by a certain normative framework, and internal when it is a psychological state of mine – when I have accepted that I have a commitment. These two uses can be applied to the terms "desires" or "wants" as well – "You should want to help your family" and "I do want to help my family".

An obligation or commitment (I use the two terms interchangeably) that I have accepted is a motivator (desire as I have defined it), as it has the same direction of fit. It has the world-to-mind direction of fit. The condition of satisfaction of an obligation, like other desires, is that it can be fulfilled or frustrated, not true or false. A commitment has the structure: I am committed to Φ (return

the wallet). It can be true or false *that* I have an obligation, but the obligation itself (internal) cannot be true or false. The same is true of other desires. An obligation can be fulfilled, like desire, just in case the world changes to meet the content of my obligation. Suppose I have an obligation to help a friend. That obligation is fulfilled if and only if I do help my friend. A state with the world-to-mind direction of fit is a motivating state as it compels us to change the world to match the content of that state.

In chapter 2, I said that a commitment is, according to Searle, the adoption of a course of action or policy (or other intentional content) where the nature of adoption gives one who adopts it a reason for pursuing the course.[152] Why is this so? There is special form of commitment, Searle argues, where one creates a commitment to another person through the "imposition of conditions of satisfaction on conditions of satisfaction".[153] The condition of satisfaction of a commitment is, as mentioned before, that it is fulfilled, that the world changes to match the content of the commitment. If one asserts that one has a commitment, one is then, by the nature of such a speech act, committing oneself to the truth of the assertion. The assertion is true only if I fulfil my commitment. This is what Searle means by claiming that we impose conditions of satisfaction on conditions of satisfaction and this is done by voluntarily asserting that I have a commitment.

This form of commitment, furthermore, is binding on the agent. Searle explains:

… the speaker stands in special relation to his own assertions, in that he has created them as his own commitments. He has freely and intentionally bound himself by undertaking his commitments. He can be indifferent to the truth of someone else's assertion, because he has not committed himself. He cannot be indifferent to the truth of his own assertions, precisely because they are his commitments.[154]

These desire-independent reasons (I am not committed to calling a reason based on the recognition of a commitment desire-independent – this is Searle's terminology relating to his thesis about assertions) motivate us in the same way as desire-dependent reasons do. To accept that I am under a commitment and that I can fulfil my commitment by Φing in circumstances C is the same as recognising that I have a desire that can be fulfilled by Φing in circumstances C. The result is the same in both cases: I am, if rational, motivated to Φ in C.

Searle's account, it can be objected, only considers speech-acts and is silent about how beliefs play a role in motivation when the belief is not expressed as an assertion. In response, look at the important similarity between an assertion and a sincerely held belief. Both involve a commitment to the truth of the proposition in question. Both impose conditions of satisfaction on conditions of satisfaction. Believing that I am morally required to Φ in C is to believe that I have an obligation to Φ

---

[152] Searle, J. p. 175.
[153] *Ibid.* p. 174.
[154] *Ibid.* p. 176.

in C. I say more about this in chapter 4. Sincerely believing a moral proposition, then, is to accept a commitment. Sincerely believing, like asserting, places the agent under a commitment. To recognise a commitment is to recognise a desire. A commitment is a desire that we can reach without any appropriately related element being present in the S.

Believing a moral proposition is different from asserting, a second objection runs, as making a speech act is voluntary in a way that believing a proposition is not. This is true, but this difference is one that a defender of moral reasons as categorical reasons for action, not his opponent, should be happy to accept. The fact that an assertion is voluntary in a way that belief is not does not undermine the crucial claim that sincere belief is, like an assertion, a commitment to the truth of the proposition in question. The involuntary nature of belief explains why a moral requirement is a requirement for all rational persons, independently of the contents of their S. The commitment to the truth of a sincere belief is why a moral obligation is binding on an agent; it is binding when the agent accepts that it is a moral obligation, that is, sincerely holds a moral belief of the form, "I ought to Φ in C". The contents of one's S are irrelevant when it comes to reasons for belief. If there do are objective moral facts (be they metaphysically or epistemically objective), and I believe there are, there is a reason for all rational persons to believe them. Sincerely believing a moral proposition does, I have argued, provide one with a reason for action. From this, it follows that moral reasons are categorical reasons.

To sum up, commitments are desires, broadly construed, as they have a world-to-mind direction of fit and a commitment to Φ in C is a desire to Φ in C. The Internalism Requirement is fulfilled, as when one recognises that one has an obligation one is motivated to act accordingly except in cases of irrationality. The belief does not motivate directly, but directly gives rise to a commitment, which is a source of motivation. This process does not require any extra source of motivation – the motivation is internal to the judgement. I will explain this point further in chapter 4.

How do we come to have commitments? I will briefly sketch an account of one way (there may be others) I think that rational agents come to have moral commitments. The first step is that the person must have a general moral belief (set of beliefs) on the basis of which the person comes to hold other more specific moral beliefs. This general moral belief can range from believing in the Categorical Imperative or the Utilitarian Principle to believing in Christian Ethics (as that person interprets them) to holding a set of moral beliefs because that was what one was brought up to believe. The point is that this belief (set of beliefs) must be able to support specific beliefs about what the person should believe to be right or wrong or how that person should act in specific circumstances. Part of holding moral beliefs is that you believe not only that some actions are right or wrong, but also that you are required to act in certain ways in certain situations. Many moral theories are explicitly stated in the form of an

action prescription. The Bible, for example, claims not only that murder is wrong, but also that "thou shalt not kill".

The next step is to look at the situation one finds oneself in to ascertain whether the specific moral beliefs one holds are relevant to this situation or not. It is at this stage that we stop reasoning theoretically and begin reasoning practically (although I believe that human reasoning normally involves a complex interweaving of theoretical and practical Reasoning. For simplicity's sake, I shall say practical reasoning starts here). Once we have properly taken account of our situation, we begin deliberating, based on what our preceding moral beliefs are, about how we believe we should act. We then make a judgement of the form, "I ought to Φ in these circumstances." Once we recognise that we have made this judgement, we recognise that we have a commitment: I am committed to Φ in these circumstances. Since a commitment is, as I have argued, a desire, the commitment plays the same role in motivating action as any other desire does that gives me a reason to act. Put more schematically, the situation I described looks as follows:

General moral belief: Do unto others as you would wish them to do unto you. (Just about any such

general moral belief will do)

Specific moral belief: Alice believes that she ought to return a lost wallet that does not belong to her

that she finds in appropriate circumstances.

Recognition of the situation: The circumstances now are appropriate for her to return the lost wallet.

Commitment: Alice ought to (has a commitment to) return the lost wallet in this situation.

This is a rough guide of how we come to hold moral commitments. The next chapter will look at this case in more detail and respond to objections that arise.

Chapter 4: Commitments, Values and Rational Motivation

In this chapter I shall show how the main argument of this thesis attempts to reconcile Ethical Cognitivism, Internalism and a correct account of Practical Reason. My account is compatible with metaphysically and epistemically objective versions of Cognitivism. All I require is that the agent, when entertaining a moral judgement, takes the judgement to be true. The route the agent takes in reaching her moral judgement (so long as she takes it to be true), although a major issue in meta-ethics, is not relevant to my project here. I aim only to show that a sincerely held moral judgement that can be true or false, however it may have been formed, does internally motivate rational persons to act according to that judgement. The judgement motivates the agent intrinsically, my account suggests, according to the version of Internalism I defended in chapter 1: Rational-Agent-Motivational-Internalism (RI): If an agent judges that she ought to perform (or refrain from performing) an action, she will be motivated to perform (or refrain from performing) that action *because* she holds that judgement, and, if she is not so motivated, she is practically irrational. The agent forms a commitment (source of motivation) *because* she holds the moral belief – nothing else is required for rational agents.

The account of Practical Reason with which I attempt to reconcile Ethical Cognitivism and RI is Humean in one way and anti-Humean in another. It is Humean because of its insistence on the presence of both a belief and a desire to motivate action, where belief and desire are defined and correspondingly distinguished according to their different directions of fit. I defined desires as states that have the world-to-mind direction of fit – states that cannot be true or false, but only fulfilled or frustrated. Beliefs, to state an uncontroversial minimal definition, are states that the agent takes to be true. Beliefs have the opposite direction of fit to desires. My account of Practical Reason is anti-Humean in that it allows, indeed requires, that it is possible and rational for one to form a desire to Φ by holding a sincere belief that one ought to Φ. One of the main aims of the present chapter is to apply my account of Practical Reason to cases of practical reasoning, which has only been provided in outline as this is all that is necessary for my purposes. To begin with I will analyse in greater detail the account I gave at the end of the previous chapter of how persons come to have a commitment.

How we form commitments

The first step is that the person starts with a general moral belief.

General moral belief: Do unto others as you would wish them to do unto you. (Just about any such general moral belief will do). The second step is that the person derives specific moral beliefs from that general moral belief.

Specific moral belief: Alice believes that she ought to return a lost wallet that does not belong to her that she finds in appropriate circumstances. The first step may not be necessary as we can imagine a person who holds moral beliefs about how she should act in certain situations without reference to one or many general guiding principles. I suspect, however, that such persons – if they are to be consistent in their moral appraisals of situations – tacitly employ some kind of general principle. Common examples are persons who form beliefs about how they should act according to how they feel about certain situations. In these cases they would be tacitly invoking a principle such as: one should avoid actions that give rise to a feeling of disapprobation and commit actions that would give one a feeling of approbation. This is roughly Hume's approach to normative ethics[155].

The reader may feel that I am moving away from Cognitivism here. The first point to note, however, is that I do not need to deny, and indeed do not want to deny, that one's moral sentiments do motivate moral action. I only need to deny that the moral sentiments are the only source capable of motivating moral action. Secondly, let us imagine one who forms moral judgements as Hume suggests and that this person is consistent in the moral judgements she makes. What else could bring about this consistency apart from a belief that I ought to avoid actions that give rise to a feeling of disapprobation and commit actions that would give one a feeling of approbation and a stable disposition to feel the appropriate sentiments in appropriate situations? I do not want to be side tracked by arguing that Hume was not a non-Cognitivist and against non-Cognitivism. The point is that I intend my account of how persons come to hold commitments and be thereby motivated to be broad enough to claim that the way that we are motivated to act according to our moral beliefs is roughly the same as the way we are motivated to act according to our moral sentiments.

It is a fact that people do hold general moral beliefs such as "do unto others as you would have them do unto you" and specific moral beliefs such as "I ought to return a lost wallet that does not belong to me that I find in appropriate circumstances". Whether these beliefs can be justified or ultimately grounded in moral facts is another matter. This is an issue that is challenged by what I referred to as "direct scepticism" in chapter 1 (p. 8). The rationality and justification of holding beliefs such as those mentioned above is of a different kind from the issue of Practical Rationality that is relevant to my argument. Practical Rationality refers to how we ought to be motivated to act or what

---

[155] Hume, D. p. 469. And p. 607 – 614.

intentions we should form relative to what beliefs and desires we happen to have or the norms that guide good practical reasoning.

What precisely is involved in an agent holding a specific moral belief such as the example mentioned above? The form of such a belief is: I ought to Φ in circumstances C. If it is a belief, the believer is committed to the truth of the proposition in question. Let us break down this belief to see precisely what the agent is committed to. "I ought" implies that the agent recognises that there is some normative requirement that applies to *her*. If the belief is a sincerely held belief or put another way, if the agent recognises that her belief is a moral belief, she will recognise certain features implicit in a "moral ought". An "ought" is an imperative and it indicates an obligation on the part of the person the imperative is directed towards. Why is it that making a moral claim on a person is to place that person under a commitment, or from a first person perspective, why is recognising a moral claim that applies to oneself the same as recognising that one has a moral commitment or obligation? To put it simply, to say of someone that they have a commitment to Φ just is to say that the person ought to Φ. The term "ought" like the term "commitment" can be used in both a moral and a non-moral sense. A commitment is a normative requirement about how I ought to act (or believe) and that normative requirement is relative to the context of the commitment. If I have a commitment to meet Brendan at the pub after we have finished our tasks, the normative framework of friends making agreements generates that commitment. If I believe that x is a desk, then I am committed to the logical consequences of that belief, one of which being not to hold the belief that x is an elephant. This commitment is generated by the normative framework governing beliefs. In this case, I ought not to believe that an object is, at the same time, two different things that are mutually exclusive. This normative framework generates the requirement of rationality that I ought not to believe that x is an elephant.

The "I ought to" in "I ought to Φ in circumstances C" is the recognition of some requirement upon me generated by a certain normative framework. In this case, because we are discussing a moral belief, the normative framework is that of morality. The requirement generated by this particular normative framework is a moral obligation or commitment. If a person professes to hold a moral belief that she ought to Φ in C and refuses to accept that she has a moral obligation to Φ in C, then that person does not sincerely hold that belief – the belief is not properly a moral belief. It would seem as if such a person does not properly understand what is involved in a "moral ought".

The next step in the process of coming to have a commitment is to recognise a relevant situation where the moral belief is applicable. Recognition of the situation: The circumstances now are

appropriate for Alice to return the lost wallet that she found. This step is a bit tricky as it happens where Theoretical and Practical Reason intertwine. Recognising that this situation is one in which she ought to return the wallet is a piece of reasoning aiming both at belief and towards figuring out how to act. Alice needs to take the relevant features of this situation into account in order to form a belief that it is appropriate to return the wallet. Alice also needs to appraise the relevant features of the situation in order to know whether she should return the wallet or not and how to do so. Even if we concede that recognising a relevant situation is theoretical rather than practical reasoning, it is integral to knowing if one should, and how to, act in a given fashion and therefore, as Murdoch insists, a key feature in deciding how to act.[156]

Once Alice has a specific moral belief about how she should act in a given situation and has recognised that the situation she is in is appropriate for her to return the wallet, she will come to form a moral judgement. In this case her moral judgement is that she ought to return the lost wallet that does not belong to her in this situation. Let us suppose that she finds the wallet on a desk in a lecture theatre and the owner's student card is inside and she is therefore able to identify the owner. The big question is how can Alice's belief that she ought to return the wallet that she finds in the lecture theatre give her a desire to do so? The desire that results from her moral belief is a commitment (see my argument in the first half of chapter 3). The big question rephrased, then, is "how can Alice's moral belief give rise to a commitment?" Williams claims that in order for a state to play a motivating role, it must be psychologically real and it must be a desire (broadly construed). I accept both of these demands and claim, further, that one can form a desire that is psychologically real by believing that one ought to act in a certain way. How can this be?

Alice's desire to Φ comes from her belief that she ought to Φ and a belief is psychologically real. I rebutted arguments that claim that one cannot come to have a desire to Φ because of a belief that one ought to Φ in chapter 2 (see p. 36 - 38). In chapter 3, I argued that when one holds a self-referential belief (I ought to Φ in C) one is, because one takes that proposition to be true, imposing conditions of satisfaction on conditions of satisfaction (p. 56 – 57). Take Alice's belief that she ought to return the lost wallet that she found in the lecture theatre, for example. Her mental state has the form "I believe that p". I will speak from a first person perspective as if I were Alice because, as I argued in chapters 1 and 2, Practical Reason should be analysed from a first person perspective. The condition of satisfaction of a belief is that it be true, in this case that p be true. That "I ought to return the lost wallet that I found in the lecture theatre" is true if there is a genuine normative requirement that I return the

---

[156] Murdoch, I. p. 421 – 435.

lost wallet that I found in the lecture theatre. For brevity's sake I use "Φ" to refer to "return the lost wallet that I found in the lecture theatre".

The proposition, "I ought to Φ" just is the proposition, "I have an obligation to Φ" or "I am obligated to Φ". There is more to the condition of satisfaction of the belief at issue than mentioned in the previous paragraph. It is a self-referential belief. This special feature of the belief adds a new dimension to the condition of satisfaction – in order for the belief to be true, I must have the commitment. It could be objected that I am equivocating on the word, "commitment". In the first usage, I mean "commitment" in an external sense – a commitment is something that is generated by a certain normative framework. The second, internal, usage of "commitment" is as a desire. My argument could only work, the objection runs, because I change the meaning of commitment during the course of the argument. In response, we must notice that when I believe that I have a commitment (external sense) to act in a certain way, this is a self-referential belief. Because I believe that the commitment applies to *me*, I attach the internal meaning of "commitment" to the external meaning and the self-referential nature of the belief means that, in taking the belief to be true, I take the commitment to apply to *me*. The key to understanding how believing that one has a commitment (external sense) leads one to believe that one has a commitment (internal sense) is, therefore, to understand the special nature of self-referential beliefs and their conditions of satisfaction.

Let me put forward what I think is the most powerful objection to my account.[157] It seems that I am agreeing with Williams and merely re-stating the terms of the debate. I am re-naming an external-reason statement an "external commitment" and I am re-naming an internal reason-statement (or the crucial part of an internal-reason statement) an "internal commitment". Through claiming that a commitment (internal) is necessary for motivation, I am only agreeing with Williams that the only reason-statements that can be correct (can motivate action) are internal reason-statements. All that my account of self-referential beliefs can accomplish (at best) is another way that we can come to have a new desire. Williams's main point still stands: only internal reasons really exist (can motivate) and these must refer to some element in one's S.

Although this objection carries some force and I do agree with Williams that desires are necessary for motivation, I disagree with him on a crucial point. This point is that I argue, *contra* Williams, that one can come to form a desire to Φ because one believes that one ought to Φ even when there is no pre-existing element in one's S related to Φing. This is crucial because it opens the way for external-reason-statements to be true. What explains one's being motivated to Φ is one's believing that

---

one ought to Φ and this is because of the nature of one's believing an external reason statement – it is a self-referential belief. We can see the explanatory role of the external reason (or reason-statement) counterfactually. Were one not to believe the external reason-statement – that is – believe that one ought to Φ in circumstances C, one would not be motivated to Φ in C. External reasons motivate action in rational persons when they are mirrored by an internal reason, and this occurs because the agent believes that the external reason applies to her.

To believe that I have a commitment to Φ is to believe that I am committed to Φ, just as my believing that I have a desire to Φ is to believe that I desire to Φ. Now, let us look at the form of the state, "I am committed to Φ". This state has the desire-like or world-to-mind direction of fit unlike my belief "that I am committed to Φ", which has the opposite direction of fit.

Does believing that I have a commitment (internal sense) translate into my having a commitment? Since it is possible that I can be deceived about what desires I have – I can believe that I desire to Φ, but upon reflection and perhaps by the exercise of the imagination I can realise that I have no such desire – I can also be deceived about what commitments I have. Where does my argument stand at the moment? I have shown that by holding a self-referential belief, one can come to believe that one has a commitment (desire) to Φ. Does this mean that I actually do have such a commitment? Does my believing that I have a desire to play golf naked in the rain mean that I actually have this desire? It is, recall chapter 2, a large part of Practical Reasoning and indeed Practical Rationality to work out what ends one actually has and to specify those ends. If this were not a norm of rationality, it would be rational to attempt to satisfy any desire that I believe that I have, even when this would thwart my other ends or when the desire I believed I had turned out, on closer inspection, not to be a desire of mine. Even if we accept some preference fulfilment account of Practical Rationality, we should accept that deliberating about one's ends is a normative requirement as this is often necessary to achieve preference fulfilment.

How should one go about determining whether a desire (commitments included) that one believes that one has is actually an end of one's? The answer, I think, can be found in Williams. One should exercise the imagination to see if one retains the desire once one realizes what its fulfilment would be like. Once one has exercised the imagination, one might find, upon introspection, that one no longer believes that one has a certain state with the world-to-mind direction of fit. Part of this process is to work out if the desire I believe that I have is consistent with certain other ends of mine (elements in my S). Williams would include one's values in one's S. In order to work out whether I actually have a certain moral commitment, it makes sense to investigate whether this moral commitment (or the

imagined fulfilment of my commitment) is consistent with my moral values. This is how I propose one determines if one has a commitment that one believes one has.

I need to introduce a caveat at this point. I do not claim that one's values are necessarily elements in one's S, as Williams does. If a value is necessarily an element in one's S and we know that we have a commitment by finding out that it is consistent with a value, this means that my commitment can be reached by rational practical deliberation starting from an element in my S. It is, however, precisely my burden in this thesis to show that rational practical deliberation need not start from some element in the agent's S (although I argue that it can give rise to a new element in one's S that was not derived from some pre-existing element in one's S). It therefore seems to be my burden to show that one's values are not necessarily elements in an agent's S, although I can allow that some persons' values are elements in their S.

Valuing

Smith argues that valuing is a mode of believing and not desiring.[158] He claims that valuing Φing is equivalent to accepting a normative reason regarding Φing – that is – accepting a rational justification of Φing or "taking" Φing to be rationally justified[159]. Smith's strategy is to argue that accepting a rational justification of Φing is a mode of believing. I think that Smith's equating valuing with rational justification without argument tilts the issue of whether or not valuing is a mode of believing or desiring in his favour. To accept that something is rationally justified has the formal features of belief. The words "to accept" function as a placeholder for a propositional state, be it a belief or desire (the word "accept" seems to fit more easily with belief than desire). I accept that p, where p is the proposition that "Φing is rationally justified." That p can be true or false, not fulfilled or frustrated – it can be true or false that Φing is rationally justified, but it does not make sense to say that "Φing is rationally justified" can be fulfilled or frustrated apart from "fulfilled" meaning to "make true". This is, of course, not how Smith argues that valuing is a mode of believing. We need, therefore, to look for a definition of valuing that does not, already, favour an account of valuing as believing.

There are different ways in which people hold values. For my purpose it is important to distinguish between a moral and a non-moral sense of valuing. I value, say, Jacques Kallis's contribution to the South African batting line up (and bowling line up for that matter). This is an example of a non-moral usage. My valuing racial equality is an example of a moral value. When enquiring whether valuing is a mode of believing or desiring we are inquiring about what it is that people do when they value something. My enquiry is, therefore, concerned with the meta-normative

---

[158] Smith, M. p. 133 – 150.
[159] *Ibid.* p. 132 - 133 and again on p. 137.

status of valuing.[160] This investigation, however, mirrors that mentioned in chapter 1 about Ethical Cognitivism. Do moral values express beliefs, constitute beliefs, or are they just beliefs, or do our moral values have some other non fact-stating role?

As I have remarked before, it is not within the scope of this thesis to argue for Ethical Cognitivism. My project is to reconcile Ethical Cognitivism (of whichever variety) with Internalism and a correct account of human motivation. I can, for the purposes of this thesis therefore, assume that one's moral valuing, at least some of the time, is a mode of believing – it can be true or false. Working from this assumption, it is possible for Alice to deliberate rationally about whether the commitment (desire) she believes that she has is consistent with her moral values and these values need not be elements in her S – they can be moral beliefs.

My account of determining if one's desires are consistent with one's values can answer the worry I recognised in footnote 122 of chapter 2: If a norm of rationality claims (roughly) that we ought to seek to maximally satisfy (or satisfice) our desires in certain situations, and this norm makes no reference to the content of one's desires, then it could be rational to maximally satisfy one's desire for murder, for example. One ought to see if one's desire(s) for murder is/are consistent with one's moral values and one would often find that they are not. If they are, however, then one ought not to have those values – should hold different moral beliefs. The claim that one ought not to value murder is a normative ethical claim. Whether or not this normative claim can be ultimately justified is a separate matter. The aforementioned worry arises, therefore, not because of my account of Practical Reason, but because of direct scepticism about the ultimate justification of morality, which is a separate matter that is beyond the scope of this thesis.

Let us return now to the process of rational practical deliberation in forming commitments. I have established that an agent who sincerely holds a moral belief will come to believe that she has a commitment to act as that belief prescribes. I claimed, and I think Williams would agree, that in order to deliberate about whether or not one actually has a certain desire (with moral content) that one believes one has, one should deliberate as to whether that desire is consistent with one's moral values or not. The desire in this case (a commitment) is to return the lost wallet that Alice found in the lecture theatre. She wants to see if this is consistent with her values. Now we have been supposing that Alice values the moral principle, "Do unto others as you would have them do unto you" and that this principle generates the specific moral belief, "I ought to return a lost wallet that I find in appropriate circumstances", which is one of Alice's specific moral values as it is derived from her general moral

---

[160] *The Encyclopedia of Philosophy.* Volume 8. 1967. Edwards, P. [Ed] Collier-MacMillan Limited. p. 231.

value. Alice will deliberate to determine whether or not her commitment to return the lost wallet she found is consistent with her values. If she finds that returning a lost wallet that she finds in the lecture theatre is consistent with her valuing the principle, "Do unto others as you would have them do unto you", she will find that her desire (commitment) that she believes she has is consistent with her values and is, therefore, actually a desire that she has. Notice that her process of rationally deliberating about whether she has the commitment she believes that she has is similar to the process that led her to form the commitment. What this hints at is that her process of coming to form the commitment was a rational process.

In order to show that Alice's process of practical deliberation to confirm that she has the commitment is a rational one, I will show how it is similar to a case that, I believe, Williams would be happy with. Suppose Cilla cares about her cat and finds out that her cat is sick. She believes that by taking her cat to the vet, the cat will be healed. She believes that she has the desire to take her cat to the vet. Cilla wants to deliberate about whether she actually has the desire to take her cat to the vet or not. She deliberates about whether or not she has this desire and tries to find out if it is consistent with her ends and values. Cilla will realise that she does care about her cat's welfare. If she finds that taking her cat to the vet is consistent with caring about her cat's well-being, she will have confirmed, through rational deliberation, that she does have the desire to take her cat to the vet. The process of rational deliberation Cilla employs to discover whether or not she has the desire she believes she has is similar to the process that led her to believe that she had this desire in the first place. This, again, implies that Cilla's process of forming her desire to take her cat to the vet was, in the first place, a rational one. The main difference between these two examples is that Cilla's rational practical deliberation started from an element in her S (caring for her cat), whereas Alice's deliberation did not start from an element in her S, but from a moral belief. But I have argued against the view that it is a necessary condition of rational practical deliberation that it must start from an element in one's S and have argued that we can come to hold a desire to Φ through holding a belief that I ought to Φ.

The Amoralist

I must now honour a promissory note by explaining the rational failure of Alex, an Amoralist. Alex believes that eating meat is morally wrong and despite this belief is in *no way* motivated to refrain from doing so. This is problematic because of the Internalism requirement, which claims that if Alex holds a moral judgement, he must be motivated to act on that judgement because of that judgement itself, or else he is practically irrational. In order to explain Alex's situation I will apply a similar analysis to the case of Alice coming to form a commitment to return a lost wallet. Alex holds a specific

moral belief, that eating meat is morally wrong in normal circumstances. I will not deal with the role of a general moral belief, in this case, because we can rest content that Alex holds a specific moral belief. Alex would then recognise a particular situation that is relevant for the application of his moral belief. In this case, he has a choice between ordering a spaghetti bolognaise or a vegetarian dish. We would then expect Alex to combine his moral belief and recognition of the relevant situation to form the judgement that he "ought not to choose the spaghetti bolognaise". Once he has made this judgement, we would expect Alex to be motivated to choose a vegetarian dish and not the spaghetti bolognaise. But the example tells us that Alex is in *no way* motivated to act on his moral belief – he is amoral. Where does Alex's failure (if it is indeed a failure) come in?

We are assuming that Alex sincerely holds the moral belief, so the failure must come in either between his holding the belief and forming the judgement, or between his holding the judgement and failing to be motivated. If it is the former, the problem is that he has not recognised the situation to be one relevant to his moral belief when, indeed, he should have. This is a failure of recognitional rationality. This failure also occurs in a case where someone fails to be motivated to adopt the acknowledged necessary means to one's end. Suppose I want to eat a meal at a restaurant and believe that in order to do so it is necessary to order my meal from a waiter. The waiter comes to my table and takes the orders of my dining companions and I fail to recognise that I should give my order to the waiter – that is – I fail to recognise that the situation is apt for me to perform the acknowledged means to my end. Because of this failure in recognising an appropriate situation, I fail to be motivated to give my order to the waiter. This failure of rationality is due to my not "getting the world right", which, as I showed in chapter 2 (p. 32 - 33 and 40 - 41), is a norm of Practical Rationality. If Alex fails to make the moral judgement that he ought not to order the spaghetti bolognaise from his moral belief, it is, similarly, a failure of recognitional rationality because he fails to "get the world right".

The second, and more interesting, explanation of Alex's failing applies when he makes the judgement that he ought not to eat the spaghetti bolognaise and fails to be motivated to do so. This occurs because Alex does not form a commitment not to eat the spaghetti bolognaise because he does not accept the commitment which his belief commits him to accepting. He does not recognise that his self-referential belief imposes conditions of satisfaction on conditions of satisfaction and that this entails that he ought to have a commitment. That this is an entailment has been established earlier in this chapter (see p. 62 – 64). In this situation, Alex has not properly formed a moral judgement because a complete moral judgement (in so far as it is distinct from a specific moral belief) requires both that he believes that he ought to Φ in the current situation and that he recognise that he has a commitment

(internal sense) to Φ in the current situation.[161] I still need to show what is wrong with Alex if he does not form a complete moral judgement from his specific moral belief. Why is failing to accept the entailments of one's beliefs a failure of *Practical Rationality*? It is a failure of Practical Rationality when it is a necessary part of successful practical reasoning.

Suppose that Alex desires to watch a movie at 20: 00 and it is 19: 30 now. Suppose, further, that Alex believes that the only way to get from his house to the cinemas is to drive directly to the cinemas along a main road and he believes that he needs to get from his house to the cinemas in half an hour. Alex, however, does not accept the entailment of his beliefs that he now needs to drive directly to the cinemas on the main road. Although this is a failure in Theoretical Rationality, this failure prevents Alex from acknowledging the necessary means to his end. If one is reluctant to concede that certain failures of Theoretical Rationality constitute failures of Practical Rationality, one ought to concede that Alex's case is still a failure of rationality. There would be no problem, if this reluctance is justified, in adapting my version of Internalism, in a non *ad hoc* fashion, to claim that a failure to be motivated by one's moral judgement is irrational – it is practically irrational when the relevant failure is in Practical Rationality and theoretically irrational if the relevant failure is in Theoretical Rationality. This modification is true to the spirit of Internalism (as I have defined it) as Rational-Agent-Motivational-Internalism's defining claim is that moral judgements motivate rational persons. A similar analysis would apply in the case where I fail to recognise that the situation is apt for me to place my order with the waiter when he comes to my table. This failure prevents me from being motivated to act on my practical judgement even if it is a failure of Theoretical, not Practical Reasoning, *contra* what I have claimed.

The final explanation of Alex's rational failing is appropriate when Alex has formed a moral judgement – he believes that he ought not to eat the spaghetti and recognises that he has a moral obligation (internal sense) not to eat the spaghetti – and is completely unmotivated to act on his moral judgement. In this case Alex recognises that he has appropriately related states with opposite directions of fit that are necessary (sufficient if he is rational) for him to be motivated. [162] This situation shares the same relevant features as Tortoise, who desires to eat chocolate and believes that by unwrapping the Bar-one in his hands and putting it into his mouth he will be able to eat it, but is completely unmotivated to unwrap the chocolate and put it into his mouth. We could also allow that Tortoise

---

[161] Nothing in my argument turns on whether or not a complete moral judgement should be described as a combination of a specific moral belief and a recognition of a commitment (internal sense) or as only a specific moral belief because I argue that one will come to have the commitment through sincerely holding a specific moral belief on pain of irrationality. Either option is compatible with my Internalism.
[162] Alex recognises a belief about how to fulfil his commitment as well.

accepts that he ought to unwrap the chocolate, but is unmotivated to do so. Once we recognise that the appropriate necessity involved in being motivated to act on one's practical judgement is a rational, not logical, necessity we can see how a failure of rationality can break the link between judgement and motivation (see chapter 1 p. 18 – 19). Alex and Tortoise are practically irrational, in this instance, because depression, spiritual or physical weakness interferes with what they have judged that they have a reason (motivating) to do.[163] Alex, similarly, can hold the practical (moral) judgement that he ought not to order the spaghetti bolognaise, but be completely unmotivated to refrain from ordering the meat dish. It is rationally, not logically, impossible and factors such as depression can interfere with one's rationality. This is a common fact of human experience. Stocker writes:

Through spiritual or physical tiredness, through accidie, through weakness of body, through illness, through general apathy, through despair, through inability to concentrate, through a feeling of uselessness or futility, and so on, one may feel less and less motivated to seek what is good [what one recognises one has a reason to do]. One's lessened desire [motivation] need not signal, much less be the product of, the fact that, or one's belief that, that there is less good to be obtained or produced as in the case of a universal Weltschmerz [world-wide weariness of life].[164]

Stocker is claiming that in cases such as these "depressions"[165] one can come to lose motivation to do what one judges to be (in one's own) good and this need not occur because the depressive prefers her depression.

Those who claim that cases such as that of Tortoise are always ones where the agent is motivated less to eat the chocolate relative to some other desire do not take instances of depression or irrationality seriously for it may be that Tortoise does not have any relevant contrary desires that provide him with a reason to act at that time. Instead of unwrapping the chocolate, Tortoise does nothing and doing nothing does not always require a desire to do nothing for it is not always an intentional action.

It can be objected that my argument showing that the Amoralist is irrational misses the point of the debate because we are really interested in explaining why the Amoralist should care about being rational.[166] If Alex's rational failing is to fail to be motivated to act on his acknowledged end (moral commitment in the internal sense), he fails to obey the means/end rule. If one does not obey the means/end rule, nothing can count as a reason at all (see chapter 2 p. 39 - 40). But to ask "why" the Amoralist should care, is to ask for reasons and this request only makes sense if the Amoralist is capable of having reasons at all. If the Amoralist's rational failing is not to form an internal

---

[163] Smith, M. p. 135.
[164] Stocker in Smith, M. p. 120.
[165] *Ibid.*
[166] I would like to thank Veli Mitova, a PhD student at Cambridge University, for bringing this point to my attention and helping me to figure out how to respond to it.

commitment from his external commitment, then he is not properly taking his belief to be true. But why should he care about this? There are two possible responses to this question. The first is that we should be doxastically responsible (in the sense of aiming to believe true propositions) because this is necessary for successful action and the second response is that aiming at truth is constitutive of what it is to hold a belief.[167] Defending either of these responses is too large a task to tackle here. We can rest content, however, with my argument that the Amoralist is irrational because this is consistent with my interpretation of Internalism. If this is, as I have argued, the case, then the Amoralist is no longer problematic for the compatibility of Ethical Cognitivism, Internalism and a correct account of motivation.

I have shown how one's moral beliefs can motivate one intrinsically, in accordance with my version of Internalism, to act on that judgement. This occurs because a moral belief directly gives rise to a commitment, which is a source of motivation. Nothing apart from the judgement itself (moral belief and commitment) or nothing that is not directly derived from the moral belief is required to motivate rational agents. I have argued that a correct account of Practical Reason does not prevent this motivation and that the process that leads one to be directly motivated to act on one's moral beliefs is a rational one. It is in this way that I have argued that Ethical Cognitivism of any form, Internalism, and an accurate account of Practical Reason are not, in fact, incompatible.

---

[167] The first response is problematic because it offers a non-epistemic justification of epistemic justification. The second response is problematic because it has to say that someone who does not properly aim at truth in their beliefs is either not holding a belief at all or is not a "good" believer. The former case renders a situation such as the Amoralist described above impossible. In the latter case, the question, "why should we care about good believing?" arises. Solving this problem is the subject of another thesis altogether and is therefore beyond the scope of this thesis.

Conclusion

The aim of this thesis has been to show how moral beliefs motivate rational persons, which has required reconciling Ethical Cognitivism, Internalism and a correct account of Practical Reason. I have attempted to fulfil this aim by arguing for a version of Internalism (RI) that allows the connection between moral judgement and motivation to be broken in cases of irrationality. This weaker version of Internalism allows me to argue that the Amoralist is indeed irrational and does not, therefore, threaten the view that moral beliefs can motivate moral action (through directly giving rise to an internal commitment) without the presence of an appropriate pre-existing element in one's S.

I claimed that moral beliefs give rise to moral commitments by arguing against the Humean view that one cannot come to form a desire to Φ simply through believing that one ought to Φ. I then argued that self-referential beliefs give rise to commitments (desires) because they impose conditions of satisfaction on conditions of satisfaction. My analysis of the concept of desire (and argument that we should understand desire in terms of direction of fit) and my analysis of belief (and sincerely held beliefs in particular) as a concept that entails certain commitments on the part of the believer, allowed me to conclude that beliefs can give rise to desires. I showed that through a process of rational practical deliberation, one can confirm that one has a moral commitment that one believes one has. Furthermore, I considered the objections to my account that I took to be the strongest and responded to these objections.

It must be acknowledged that I did not argue for Ethical Cognitivism and did not respond to potential Non-Cognitivist arguments that would be fatal to my central argument. Although this is a gap in my account, it is an immense task that could hopefully be the subject of a later work. It should also be conceded that I did not fully explain why the Amoralist should care about being rational in a case where he does not properly take his beliefs to be true and/or does not accept the logical entailments of his beliefs. This issue is, again, the subject of a separate thesis that is beyond the scope of this project.

Within the parameters of the debate that I set up, which I believe to be as broad as can reasonably be expected, I have provided an interesting argument, which concludes that moral beliefs do motivate rational persons according to the criteria set up in the Internalism Requirement that I argued for.

Bibliography:

Aristotle. *Nichomachean Ethics.* 2000. Crisp, R. [Ed]. Cambridge University Press.

Audi, R. "Moral Judgement and Reason for Action" in *Ethics and Practical Reason.* Cullity, G. and Gaut, B. [Eds]. 1997. Clarendon Press.

Brandom, R. "Actions, Norms, and Practical Reasoning". 2001. In *Varieties of Practical Reasoning.* Millgram, E. [Ed]. MIT Press.

Bratman, M. "Taking Plans Seriously." 2001. In *Varieties of Practical Reasoning.* Millgram, E. [Ed]. MIT Press.

Brink, D. *Moral Realism and the Foundations of Ethics.* 1989. Cambridge University Press.

Cullity, G. "Practical Theory". In *Ethics and Practical Reason.* Cullity, G. and Gaut, B. (Eds). 1997. Clarendon Press.

Davidson, D. "Actions, Reasons and Causes". In *Essays on Actions and Events.* 1980. Clarendon Press.

Dreier, J. "Humean Doubts about Categorical Imperatives". 2001. In *Varieties of Practical Reasoning.* Millgram, E. [Ed]. MIT Press.

Fehige, C. "Instrumentalism." 2001. In *Varieties of Practical Reasoning.* Millgram, E. [Ed]. MIT Press.

The Encyclopedia of Philosophy. Volume 8. 1967. Edwards, P. [Ed]. Collier-MacMillan Limited.

Hooker, B. "Williams' Argument against External Reasons." In *Varieties of Practical Reasoning.* Millgram, E. [Ed]. MIT Press.

Hume, D. *A Treatise of Human Nature.* Selby-Bigge, L.A. [Ed]. 1978. Oxford University Press.

Kant, I. *Groundwork for the Metaphysics of Morals.* 2002. Hill, T. And Zweig, A. [Eds] Oxford University Press.

Kolnai, A. "Deliberation is of Ends." 2001. In *Varieties of Practical Reasoning.* Millgram, E. [Ed]. MIT Press.

Korsgaard, C. "Skepticism about Practical Reason." 1986. In *The Journal of Philosophy.* Volume LXXXIII, No. 1.

Korsgaard, C. "The Normativity of Instrumental Reason". In *Ethics and Practical Reason.* Cullity, G. and Gaut, B. [Eds]. 1997. Clarendon Press.

Korsgaard, C. *The Sources of Normativity.* 1996. Cambridge University Press.

Mackie, J. *Ethics: Inventing Right and Wrong.* 1977. Penguin Books.

Mackie, J. *Hume's Moral Theory.* 1980. Routledge and Kegan Paul.

Miller, A. *An Introduction to Contemporary Metaethics.* 2003. Polity Press.

Millgram, E. "Introduction". 2001. In *Varieties of Practical Reasoning*. Millgram, E. [Ed]. MIT Press.

Murdoch, I. "The Idea of Perfection." 2001. In *Varieties of Practical Reasoning*. Millgram, E. [Ed]. MIT Press.

Nagel, T. *The Possibility of Altruism*. 1970. Clarendon Press

O'Neill, O. "Consistency in Action." 2001. In *Varieties of Practical Reasoning*. Millgram, E. [Ed]. MIT Press.

Plato. *Republic*. Translated by Robin Waterfield .1994. Oxford University Press.

Rachels, J. *Ethical Theory 1*._1998. Oxford University Press.

Railton, P. "On the Hypothetical and Non-Hypothetical in Reasoning". In *Ethics and Practical Reason*. Cullity, G. and Gaut, B. (Eds). 1997. Clarendon Press.

Routledge Encyclopedia of Philosophy. Volume 7. Craig, E [Ed]. 1998. Routledge Press.

Schmidtz, D. "Choosing Ends." 2001. In *Varieties of Practical Reasoning*. MIT Press.

Schueler, G. *Desire: Its Role in Practical Reason and the Explanation of Action*. 1995. MIT Press.

Searle, J. *Rationality in Action*. 2001. MIT Press.

Shafer-Landau, R. "Moral Judgement and Moral Motivation." 1998. In The Philosophical Quarterly. Volume 48, No. 192.

Slote, M. "Moderation and Satisficing." 2001. In *Varieties of Practical Reasoning*. Millgram, E. [Ed]. MIT Press.

Smith, M. *The Moral Problem*. 1994 Blackwell Publishers LTD.

Snare, F. *Morals, Motivation and Convention: Hume's Influential Doctrines*. 1991. Cambridge University Press.

Thagard, P. "How to Make Decisions: Coherence, Emotion, and Practical Inference." 2001. In *Varieties of Practical Reasoning*. Millgram, E. [Ed]. MIT Press.

The Encyclopedia of Philosophy. Volume 8. 1967. Edwards, P. [Ed] Collier-MacMillan Limited.

Wiggins, D. "Deliberation and Practical Reason". 2001. In *Varieties of Practical Reasoning*. Millgram, E. [Ed]. MIT Press.

Williams, B. "Internal and External Reasons". In *Moral Luck*. 1981. Cambridge University Press.

Williams, B. *Morality*. 1972. Penguin Books.