Philosophy

MORAL OBJECTIVITY AND THE PSYCHOLOGY OF MOTIVATION

By

Jude Ndubuisi Edeh

A DISSERTATION

Submitted in Partial Fulfilment of the

Requirements for the Degree of

Doctor of Philosophy (Dr.phil)

University of Münster

Münster, Germany

2017

ACKNOWLEDGMENTS

ABSTRACT

This dissertation provides a solution to the tension of specifying and reconciling the relationship between moral judgement and motivation. I defend motivational externalism against realist motivational internalism, the view that moral belief is not only a genuine belief, but also, it necessarily motivates. I argue that by making moral motivation analytically dependent on moral judgement, the realist internalist removes the moral person from the landscape of moral agency. Precisely, the conceptual intuition on which realist internalism rests is insufficient. For it saddles the burden of explaining away, for example, the crucial roles of agents' desires, self, traits and identity in moral motivation. To fill this gap, I develop a realist externalist account of moral motivation. My core argument is that motives for acting on moral judgements are based on independent desires, and the desire for being faithful to the moral-self is one of them. Given that morality has two aspects, namely theoretical feature and practical feature, I proceed by arguing that noncognitivism; precisely moral expressivism fails to capture our intuition about the theoretical feature of morality as well as moral objectivity. Then, I characterize and defend the version of moral realism relevant for the purposes of this work. Having taken a metaethical standpoint, I spend the rest of the dissertation addressing the practical feature of morality. First, I argue that the phenomenon of moral indifference threatens the motivational efficacy claim of robust internalism. In addition, I contend that Smith's Arguments from Rationalism and Practicality Requirement do not give us sufficient grounds to abandon motivational externalism. Second, I argue that motivational externalism has better resources for analysing the psychology of moral motivation. To defend motivational externalism, (1) I provide a conceptual framework; and (2) drawing on Harry Frankfurt's "Second-Order Volitions" and Augusto Blasi's "Moral-Self", I propose an externalist account of a widespread connection between moral judgement and motivation. On this view, motivation is based on moral identity. I argue that *moral care*, *moral-self* and *moral desires* are essential to moral motivation. Moral motivation does not stem directly from moral judgements; rather from an independent desire, namely, a moral desire for realizing the moral values we care about, that is, the desire to be faithful to one's moral-self. Finally, I argue that externalism is not condemned to explaining the motivation of virtuous persons in terms of desire *de dicto* as Smith alleges. In fact, it is possible, on the externalist construal, to account for a reliable connection between moral judgement and motivation in terms of *de re* desires.

Moral Objectivity and the Psychology of Motivation

Die Objektivität der Moral und die psychologische Frage nach der Motivation zu moralischem Handeln

Zusammenfassung

Die Arbeit widmet sich einer zentralen und umstrittenen Frage innerhalb der Metaethik und Moralpsychologie: Wie kann die Vorstellung, dass unsere moralischen Urteile wahrheitsfähige Aussagen darstellen, mit dem Anspruch vereinbart werden, dass diese Urteile uns zugleich zu einem entsprechenden Handeln zu motivieren vermögen? Um dieses Spannungsverhältnis zwischen moralischem Urteil und Motivation aufzulösen, verteidige ich eine externalistische Variante moralischer Motivation gegenüber einem motivationalen Internalismus. Die Analyse der Motivation zu moralischem Handeln erfolgt zwar in beiden Fällen auf der Basis eines moralischen Realismus, d.h. der Vorstellung, dass moralische Urteile wahrheitsfähig sind. Im kritischen Durchgang durch einschlägige internalistische Positionen argumentiere ich jedoch dafür, dass diese nicht überzeugend zeigen können, wie wahrheitsfähige moralische Urteile uns zugleich zu einem entsprechenden Handeln zu motivieren vermögen. Der Kern meiner Kritik besteht darin, dass der realistische Internalismus die Bedeutung der Person in der Analyse moralischer Motivation verkennt, d.h. die Bedeutung der Rolle von Wünschen und eines praktischen Selbstverständnisses, d.h. der Identität des Akteurs. Bezieht man dies hingegen mit ein, so ergibt sich ein realistischer Externalismus mit Blick auf die Motivationskraft moralischer Urteile. Diese bleiben also auf den von ihnen unabhängigen Wunsch des Akteurs angewiesen, sich in einem umfassenden Sinne als moralische Person zu verstehen und diesem Selbstverständnis im Handeln treu bleiben zu wollen. Erst vermittelt durch ein solches praktisches Selbstverständnis vermögen uns moralische Urteile folglich zu einem moralischen Handeln zu motivieren.

CONTENT

CONTENT

PART TWO: Psychology of moral motivation

CHAPTER 4: MOTIVATIONAL INTERNALISM

CHAPTER 5: THE MORAL PROBLEM: SMITH'S SOLUTION?

CHAPTER 6: OBJECTIONS AGAINST EXTERNALISM AND REPLIES

CHAPTER 7: A Defence OF EXTERNALIST MORAL MOTIVATION

Chapter 1

The Problem of Moral Motivation

"I'm good!

Why, then be moral?

Do you mean, I've reasons to?

I see it not,

I don't see why I've to be moral.

I'm good!"

Culled from Olivia, unpublished Poem

1.0.   Introduction

There is a belief about the practical implications of moral thoughts and language: an idea based on our moral intuition. It is believed that morality has two basic features, namely, the theoretical and practical features (Brink 1989; Smith 1994; McNaughton 1988; Stevenson 1937; Nagel 1970; McDowell 1978, 1979; Bromwich 2010, 2013; Dancy 1993; van Roojen 2002; Radcliffe 1999; Railton 2006). However, there is a tension between conceptions about these features of morality;[1] and it [this tension] arises from the fact that the features "pull against each other, so threatening to make the very idea of morality incoherent" (Smith 1994, p. 4). Hence, if morality is to be taken seriously as a coherent domain of human inquiry, it has to justify the nature of these features as well as

---

[1]It is important to make some remarks on how 'ethics' and 'morality' and their cognates, such as 'ethical' or 'moral', are understood here. Matthew H. Kramer in *Moral Realism as a Moral Doctrine* makes the following remarks on this issue: "The former term is more capacious than the latter. In other words, "[…] the domain of ethics encompasses the domain of morality, but also extends more widely. All moral propositions are ethical propositions, but not vice versa" (Kramer 2009, p. 2). Further, ethics in a general sense can mean a domain that deals with the issues of good life. However, it can as well mean 'theory of morals' or 'Moraltheorie'- that is, principles of right and wrong actions (for details on this, see the discussion on normative ethics in section 1.3). Although this distinction seems to hold, it becomes superficial if we construe these terms, that is, ethics (ethical) and morality (moral) in terms of relation to what is right and wrong, good and bad. Hence, for the sake of uniformity, I will try not to use them interchangeably. I will keep to 'moral' and 'morality'; and whenever 'ethics' or 'ethical' is used, a clarification will be given especially in cases where it is applied in a certain classified sense.

their relationship. The relationship in question manifests itself properly in the connection between moral judgements and motivation. However, the connection is highly debated. On the one hand, the internalists argue that moral judgment, for example, '*'p'* is morally right' is necessarily connected to being motivated to '*p'*. On the other hand, the externalists reject this claim by arguing that moral motivation is constituted by independent desires. Thus, the precise task that lies before moral philosophers, especially the metaethicists and moral psychologists, is the one of specifying and reconciling moral judgement and motivation without trivializing the connection. Nonetheless, the controversy surrounding the various understandings of this connection makes this task a difficult one. Let me start by explaining the features of morality.[2]

Two Cases (Dave, and then Olivia)

> *Dave:* Dave and his wife Lotta have four children. Dave works for a bank. He just received the news about the bank's bankruptcy decision. Apart from the fact that he was supposed to be promoted that month, he was worried about his family. Throughout the day, he thought through what losing his job would mean for his family. No sooner had he returned home than he explained everything to Lotta. Dave is worried about Lotta's chronic medical condition; he is worried about the family's health insurance coverage, losing his house and not being able to provide for his children, especially their education. However, before receiving the bankruptcy news, Dave discovered a dormant account with a huge sum of money. Upon an independent investigation, he found out that the account belonged to a customer, who died 10 years ago in a plane crash. Actually, the whole family died in that accident; hence, there is no next of kin. The bank does not know about the crash and the dormant account. In other words, Dave could make away with the money without traces. He *feels* it is the right thing to do, at least, he can use the money responsibly: with it, he can provide for his family and even donate a part

---

[2] Smith refers to these features as the objectivity and practicality of moral judgements (Smith 1994, p. 6), whereas Brink calls them the intellectual and practical dimensions of morality (Brink 1997, p. 4).

of it to an orphanage. Lotta, nonetheless, thinks and strongly believes that taking the money is wrong. She, in fact, calls it stealing. After a long argument, deliberation, and disagreement, Dave saw reasons with Lotta. They agreed that it is stealing; and, as such morally wrong.

Morality is an essential part of our lives. Every day we are confronted with moral issues and decisions; and they engage us in various degrees of complexities and commitments. We are confronted with both simple and difficult [moral] situations. Sometimes, we are presented with moral situations that demand immediate decisions and actions. At other times, we engage in long arguments and deliberations as in the case of Dave and Lotta. We argue and deliberate about moral issues because we believe that there are correct answers to our questions. The disagreement and argument that ensued between Dave and his wife, and the subsequent conclusion presuppose some sort of correctness in morality. In addition, the notion of *correctness* underlies our intuition about the theoretical feature of morality.

1.1. Theoretical Feature of Morality

In his book, *The Moral Problem*, Michael Smith says that "by engaging in moral practice [that] the participants are concerned to get the answers to moral questions *right*" (Smith 1994, p. 5). Nevertheless, due to the normative character of morality, it is debated whether there are determinate ways of assigning answers to moral questions. In other words, one might argue that since morality deals with values: 'Are there correct answers or series of possible answers?' If there are correct answers, are they dependent on agents' beliefs, conditions, and circumstances or not? Going back to Dave's example, suppose he maintains his initial position: Assuming his action will not harm anyone. Further, suppose he uses the money not just for his family needs, but also for the orphanage. Since Dave seems to have *reasons* that count in his favour, one may correctly ask: In what way, if at all, can participants separate *right or correct* from the *wrong or incorrect* answers?

Moral questions and problems generate series of possible answers, which seem to offer agents *reasons* and the possibilities of choosing answers that best suit their situations, beliefs, conditions, attitudes, opinions, etc. Assuming this idea sticks, does it

mean every answer is correct? Hence, "there would be no answer would be *determinately* correct" (Kramer 2009, p. 87). Nevertheless, the idea of indeterminacy would place morality under a serious threat. First, if every possible answer would be correct in a given issue, there will be no need then of conceiving the issue as a (moral) question in the first place. In other words, if there were no idea of correct answers at all, disagreement will then not arise (and will not be required) even in the ordinary discourse. Second, if there were no idea of correct answers, and we happen to find ourselves in moral disagreement or dispute, there will be no way of resolving the issue. Precisely, the idea of construing issues as questions or problems presupposes some sort of correctness; and this applies to the moral domain equally, because when we engage in moral disagreement we strive to arrive at correct answers. According to David O. Brink,

> Moral argument and deliberation presuppose not only correct answers to moral questions but also answers whose correctness is independent of our moral beliefs. In moral deliberation and argument we try and hope to *arrive* at the correct answer, that is, at the answer that is correct prior to, and independently of, our coming upon it. (Brink 1989, p. 31)

Notice that the notion of correctness in morality, on Brink's construal, is not based on the agent's beliefs or dependent on constructed procedures. It is, rather, embedded on moral reality – moral facts. Applied to Dave's case, it could be said that before they came to agree that 'stealing is morally wrong', the correctness of their conclusion is determined, and independent of whatever they thought, felt or said about it. It is not their thoughts and talks about the answer that made it correct. If Brink is right, then the correctness of their answer is prior to and independent of their coming upon it. Brink's position is anticipated in Nagel's *The Limits of Objectivity*:

> The ordinary process of deliberation, aimed at finding out what I have reason to do, assumes that the question has an answer. And in difficult cases especially, deliberation is often accompanied by the belief that I may not *arrive* at that answer. I do not assume that the correct answer is just whatever will result or has resulted from the consistent application of deliberative methods – even assuming perfect information about the facts. In deliberation, we are trying to arrive at

4

conclusions that are correct in virtue of something *independent* of our arriving at them. (Nagel 1979, p. 100)

This is one of the core ideas driving moral realism. Moral realists largely and confidently believe that our moral thoughts and talks are morally mind-independent[3] (Moore 1903; Shafer-Landau 2003; Scanlon 2014; Boyd 1988; Railton 1986; Sturgeon 2006, Smith 1994; Kramer 2009). Mind-independency, on the construal of moral realism, is rooted in moral facts.[4] The facts of morality, in turn, are seen as the basis of moral correctness. They set, if only roughly, the principles for determining and discovering correct answers to moral questions. Thus, the correctness of moral answer is determined by its ability to capture moral facts, independent of what anyone believes, thinks, feels or says. Coming back to our question about separating correct answers from the incorrect ones, on the moral realist's analysis, moral participants can, thus, separate the right and correct answers from series of other possible answers by considering the answer that corresponds to the objective moral facts.

Furthermore, the idea of moral correctness rooted in moral facts naturally flows from the common intuition underlying other factual domains of human inquiry. Consider the following example, when we ask, 'What is nicotine?' Let us say, we have two groups of scientists, SG1, and SG2; and they report the following:

SG1: Nicotine is hygroscopic, a colourless oily alkaloid with 162.231556g/mol.

SG2: Nicotine is an oily substance with 5462.231556g/mol.

Can we discover the correct answer? Yes, we can, namely, by considering which of them rightly describes the fact in question. If the answer of SG1 is true, then it cannot be the case that both answers are true; hence, SG2 must be false. The correctness of SG1's answer does not depend on what they think or believe about nicotine, rather on an

---

[3] However, it is important to note that the idea of mind-independence is not restricted to moral realism. We shall see in the next chapter that moral theories such as moral constructivism or procedural realism defend the notion of moral mind-independence (for example, Korsgaard 1996, 2008). There is, as well, a move in the non-cognitivist tradition which argues that expressivism is compatible with mind-independence (for example, Sinclair 2008; Gibbard 2003; Blackburn 1981).

[4] We shall return to the difference between moderate or procedural realism and robust realism in the next chapter. Both positions espouse the idea of correctness in morality; however, they differ with regards to the nature and existential status of moral facts.

objective fact. In a similar way, it is argued that the concern for correct answers to moral questions presupposes an existence of an objective reality. Smith writes that "the natural interpretation of that presupposition is that there exists a domain of moral facts" (Smith 1994, p.5). On this analysis, moral questions (for example, 'Is stealing morally right?') have correct answers and as such, we can discover them by considering the answer that tersely corresponds to the moral state of affairs independent of what we think and believe about them. Based on these moral facts, we can determine our moral answers as well as form our beliefs about them – moral beliefs. According to Smith, these are "facts about which we can form beliefs and about which we may be mistaken" (Smith 1994, p.5).

Smith further argues that the prevalence of moral discourse, especially, moral questions and problems that characterize our daily lives point to an important fact, namely, the *epistemic accessibility* and *availability* of moral facts. He maintains that "the way in which we conduct ourselves in living the moral life seems to presuppose that these facts are in principle *available to all*; that no one in particular is better placed to discover them than anyone else" (Smith 1994, p.5 emphasis is mine). In other words, if moral facts were not accessible to all (that is, to 'normal' moral participants to some extent), then the moral domain would lack one of the basic requirements of human inquiry: namely the existence of common ground facts required in any credible domain of human inquiry. Such facts give the members of a given domain-community a common platform for engaging in meaningful arguments, deliberations, talks, etc. This holds also for the moral domain, for even "when we are uncertain about moral issues, we often deliberate as if there were a right answer to the issue before us" (Brink 1989, p. 27). This attitude presupposes the existence of common facts guiding our deliberation, and such facts seem to be available to us as well.

An alternative way of explaining the intuition underlying the theoretical feature of morality is by asking whether there are correct answers to questions like: 'Is abortion or assisted killing morally permissible?' 'Is religious or sexual discrimination morally wrong?' 'Is racism ever morally good?' While it is true that these sorts of questions occupy a significant part of our lives, one might still argue that there are no correct ways

of answering or at least, resolving them. Against this latter concern, Renford Bambrough elegantly argues,

> However long and violent a dispute may be, and however, few or many heads may be counted on this side or on that, it remains possible that one party to the dispute is right and the others wrong. Galileo was right when he contradicted the cardinals; and so was Wilberforce when he rebuked the slave-owners. (Bambrough 1979, pp. 18 – 19)

Kramer as well strongly believes that there are unique correct answers to moral questions (such as the ones raised above). He argues that "any moral principle that generates a negative answer to either of those questions is incorrect for that very reason" (Kremer 2009, p. 42). We can conclude that Lotta and Dave's conclusion (stealing is morally wrong) – irrespective of their feelings, beliefs, family circumstances – illustrates the place of determinate correctness in morality. This belief drives our intuition about theoretical feature of morality. Let us then turn to the practical feature of morality.

1.2. Practical Feature of Morality

Olivia's Case

Olivia: Olivia is a professor of moral psychology. Recently, she had a long conversation with Emma, a doctoral student at the department she was visiting. They discussed the dangerous impacts of climate change, especially on women and children from poor countries. Olivia argues convincingly that we are obliged to protect our environment. She strongly believes that any action with harmful impacts on the environment is morally impermissible. At the end of the conversation, Emma came to share her moral conviction: *Environmental harm is morally unjust*. Later, Emma and some of her friends filed a petition against the university authority on the grounds of some of its environmental unethical practices. Their target was to collect 1000 signatures. Within the space of three weeks, the petition gained an overwhelming support from both the professors and students, thanks to Olivia's moral conviction. However, when Emma approached Olivia to get her signature on the petition, she *declined* to sign it.

Assuming Olivia's signature was the very last signature required or that Emma and her friends have just a few more hours to submit their petition. Assuming they were confident that Olivia would surely sign the petition given her positive judgement on the issue. Apart from their feeling of disappointment, say, for not submitting the petition as planned, it could be argued that there is something very disturbing about Olivia's behaviour. How is it possible that her moral conviction failed to make her act accordingly? What might have been the cause of Olivia's unusual behaviour? We might describe her attitude as strange, because it is largely believed that moral beliefs or convictions motivate us to act (at least to some extent). Why did her moral belief (which supposedly made a claim not only on her but also on other people) unable to move her to act accordingly? Was she insincere? Is she suffering from some sort of conditions? In any case, Smith describes behaviours such as Olivia's as a real puzzlement (Smith 1994, p. 7). The practical feature of morality is based on the intuition that moral judgements, beliefs, considerations or convictions have motivational character. However, Olivia's case is disturbing as well as puzzling because it threatens this intuition.

1.3.   Olivia's Case is Metaethical

The main worry here is as follows, if cases such as Olivia's are possible, then it will be possible for the two features of morality to come apart in *actual* as well as *conceptual* contexts. Moreover, as noted above, the coherence of morality will be threatened if these features fall apart. This raises the concern of understanding the relational status of the theoretical and practical features of morality. Precisely, it calls for studying the nature of moral judgement as well as the psychological states responsible for moral actions. Notice that these subject-matters fall under the domain of metaethics. However, before we go any further, some remarks about the branches of ethics are in order.

First, **metaethics** studies the nature of morality by inquiring into the nature as well as the status of moral properties, facts, truths, relations, thoughts, talks and moral judgements. Its inquiry occurs within the scopes of metaphysics, epistemology, and semantics of morality. It asks questions such as: Are there such things as moral facts; and if they exist, what defining characters do they possess? Are there moral truth-makers?

Are moral statements propositional and truth-apt? Can we robustly or minimally assign truth-values to moral judgements? More so, a metaethical inquiry is closely tied to psychology and action theory. On this level, it concerns itself with the explanation of the connection between, say, moral considerations, motivations, and actions – **Moral Psychology**. It is also good to note that metaethics is dubbed 'second-order' account of ethics because it deals with "issues *about*, rather than *within*, morality" (Brink 1989, p. 1; *emphasis is mine*).

Second, **normative ethics** or 'first-order account of ethics', deals with how we should act. That is, it asks questions about the content of ethics. For example, it concerns itself with issues about the principles of right and wrong actions. It provides moral norms (hence, normative ethics) that guide how we ought to decide and act in moral situations. The first-order moral judgements are "of various degrees of generality, as to what things, acts, or qualities are good or bad, right or wrong" (Gewirth 1960, p. 187). By asking how we should act, the interest is not merely on concrete norms, but also on theories that underlie our moral conduct. To illustrate, moral theories offer us some sort of general abstract principles that guide our actions. For example, if Dave were to argue that his action would produce happiness for more people, then we could, roughly, dub him a consequentialist, hence, consequentialism.[5] In other words, within the normative ethical discipline, there are concrete and abstract (or theoretical) moral questions respectively. Brink has this to say about them:

> Issues of moral theory or principle concern the theoretical structure of morality. How are goodness and rightness related? In what way should a moral theory be impartial among people? Does impartiality require us to maximise aggregate welfare, or would maximizing aggregate welfare ignore important distributional

---

[5]Consequentialism maintains that moral actions depend on actions' consequence or outcomes. In other words, if an action produces the best (or least a good) outcome, such an action is regarded as a morally right action; otherwise it is a morally wrong action (for more discussion see Sinnott-Armstrong, Consequentialism, The Stanford Encyclopedia of Philosophy, Winter 2015 Edition; Howard-Snyder, The Heart of Consequentialism, Philosophical Studies, 1994, 76, pp. 107–29; Flew 1979, *Consequentialism, A Dictionary of Philosophy*, 1979). We can refer to Lotta as a deontologist: Unlike consequentialism, deontology or duty-based ethics holds that moral actions depend on adherence to duty(ies) or rule(s) (for more discussion see Alexander and Moore, Deontological Ethics, The Stanford Encyclopedia of Philosophy, 2015 Edition).

considerations? […] Substantive moral questions, on the other hand, raise more specific moral issues. Is abortion ever justified and, if so, under what conditions? (Brink 1989, pp. 1 – 2)

To describe metaethics in more detail against the background of normative ethics, let us approach Dave's case from both normative and metaethical perspectives. The former asks, 'Is stealing morally wrong?' Whereas in the latter it is asked, 'What does it mean to say that stealing is *morally wrong*?' One of the main objectives of normative ethics is to evaluate moral statements against certain standards or norms of moral conducts; and such standards could be based, for example, on *duty*. When applied to Dave's case, taking the money would be morally wrong because it is a breach of duty. From the metaethical perspective, assuming we tie moral goodness merely to duty, one might ask, 'Is doing one's duty always morally good? Suppose Lotta's argument was based on duty, it could still be asked 'whether fulfilling one's duty, say X, is good'. Unfortunately, duty alone does not tell us everything about goodness, and as such, the list is still open. After all, Dave can still argue that his action will *produce happiness* and favour not only his family but the kids in the orphanage as well. G. E. Moore argues that defining moral terms such as goodness in terms of another property or sets of properties amounts to closing down by definition certain questions, which should remain open.[6] Alternatively put, normative ethics is not in the business of telling us what goodness means. It is, rather, the business of metaethics, because it deals with questions such as 'Is there a property called goodness at all?' 'If yes, what does it *mean*?' What do Lotta and Dave mean when they say that *something* is morally wrong? Does any psychological state support their respective statements? If only roughly, it could be said that metaethics provides the rational and semantic foundations for moral inquiry. According to Smith:

---

[6] Moore clearly believes that any attempt of defining 'good' in terms of something else opens up questions; this is because "whatever definition be offered, it may be always asked, with significance, of the complex so defined, whether it is itself good". The objective of open question argument is to refute the ontological claims of ethical naturalism. Moore argues that "good does not, by definition, mean anything that is natural; and it is therefore always an open question whether anything that is natural is good" (Moore 1903, §27). Further, he argues that defining or analyzing moral terms such as such goodness in natural terms leads to naturalistic fallacy (Moore 1903, §12).

Philosophers have surely been right to give meta-ethical questions a certain priority over questions in normative ethics. If moral argument is not simply a species of rational argument, then that calls into question the very role of moral argument in everyday life. (Smith 1994, p. 3)

Principally, Olivia's case does not fall under the scope of normative ethics: it does not ask whether her action is morally wrong. It is, rather, interested in finding out why her moral conviction was unable to motivate her to act accordingly. Her case, thus, falls under the scope of metaethics. Thus, Oliva's case is metaethical. Her case, precisely, falls under the psychological debate of metaethics, and it is within this debate that this dissertation is located. Therefore, the main question motivating the present work is: **'In what sense, if at all, can objective moral judgement motivate agents to act morally?'**

1.4.   From the Mind to Moral Psychology

The theoretical and practical features of morality imply, among other things, that by holding a moral judgement, agents are furnished with *reasons* for action, which in turn, motivate them to act (at least, to some extent). Notice that this intuition closely ties metaethics to human psychology and actions (Pigden 2007, p. 200). In other words, there is a psychology that supports the connection between moral judgement, motivation, and action. In the recent years, David Hume's contribution in this regard has continued to receive an increasing attention from scholars (Coleman 2008; Schroeder 2007; van Roojen 1995, 2002; Smith 1994, 1987; Wedgwood 1995; Cohon 1994; Baier 1991; Pigden 2007; Bricke 1996). For instance, Smith argues that we owe the standard human psychology driving the moral motivational problem to Hume (Smith 1994, p. 7). In this section, I start by laying out how Hume derived his human psychology from the philosophy of mind; and then in the next section, I show how Hume's psychology threatens realist motivational internalism (1.6).

Hume grounds his philosophy of mind on the notion of perception. According to him, "nothing is ever present to the mind but its perceptions; and ... all the actions of

seeing, hearing, judging, loving, hating and thinking, fall under this denomination" (THN 456/293). Hume places all actions – insofar as they are mental events – under perception. That is, everything, which is in the mind, is nothing but perceptions. Notice that if there is anything, *possible* or *actual*, that is in the mind, it must be perception, "whether we employ our senses, or are actuated with passion, or exercise our thought and reflection" (THN 647/408). He further divides the perceptual contents of the mind into *impressions* and *ideas*. The former are the immediate and vivid products of sensory experience; however, impressions can as well arise through spontaneous reflections. Hume also argues that passions fall under impressions. The latter is faint copies of original impressions. For example, the direct and immediate experience of snow is different from the thought of snow in summer. On Hume's analysis, ideas are dependent on impressions. Nevertheless, impressions and ideas are distinct mental contents.

Hume argues that ideas are connectable, that is, we can fuse different ideas together, thereby having an association (of ideas). We achieve this fusion of ideas through resemblance, contiguity, or causality; and beliefs, then, emerge from the recurring applications of mental associations. Beliefs can either be formed on associations of analytic ideas or experiential ideas.[7] In the former, namely, the **relations of ideas,** their associations are based on the nature of conceptual and abstract contents. For example, we can show the agreement and disagreement of arithmetic, geometrical or logical ideas through demonstration or deduction. Whereas in the latter, **matter of facts,** their associations are grounded on experience or contingent facts. Unlike analytic or conceptual belief, contingent beliefs are not based on *a priori* reasoning, but rather on causal inference or a posteriori reasoning. In both cases, the main objective is to establish the relations of agreement or disagreement, hence, truth or falsehood. The truth-value of beliefs in the relations of ideas is derived from concepts contained in the statement. For example, we can deductively prove that triangle has three angles or that a bachelor is an unmarried man. Whereas truth-value of beliefs in the matter of facts is justified through

---

[7] However, according to Hume, beliefs are not ideas, because we can have the same object of conception, while believing differently. To illustrate, let us say that F is an object of conception. A believes F, whereas B disbelieves F.

experiencing things as they are in the world. For example, recall the nicotine example, we can inductively infer whether nicotine is hygroscopic [a colourless oily alkaloid with 162.231556g/mol] is true or false by observing its properties. Notice that passions and beliefs come under separate mental categories; and as such, they differ in a number of significant ways. Beliefs are representational states: They represent things as they are in the world. Smith writes that:

> Beliefs are states that purport to represent the way the world is. Since our beliefs
> purport to represent the world, they are assessable in terms of truth and falsehood,
> depending on whether or not they succeed in representing the world to be the way
> it really is. (Smith 1992, p. 330)

If Hume is correct in claiming that "Belief is nothing but a more vivid, lively, forcible, firm, steady conception of an object"[8], then it might be the case that the success or failure of beliefs hitting their truth-target depends on the strength or force with which the objects are conceived. In any case, beliefs are characterized, on the one hand, by their representationality, propositionality, and truth-aptness. Passions, on the other hand, do not arise from relations of ideas; hence, they are different mental states.

Hume argues that these states are not products of reason. That is, they are some sort of non-representational states. Hume couches these states under the general term 'passions'. Although there are other passions such as benevolence, hope, grief, fear, I will be using *desires* as an instantiation of these non-representational states. My reason for this is the role of "desire-belief" pairing in the contemporary motivational debates (Radcliffe 1999, p. 102; Smith 1994, pp. 8 – 11). Desires, unlike beliefs, are attitudes that do not aim at representing things in the world as they are. They are not propositions, which admit truth or falsehood. Imagine, for example, Robert saying, "I want chocolate". This statement can neither be demonstrated deductively nor inductively, nor be said to be

---

[8] However, as we already mentioned, although ideas are brought together through association, it is good to note that each idea is distinct, so also, beliefs are separable from ideas and imaginations. For example, a mere conception of snow or imagination of snow differs from believing that 'Snow is white'. He writes, "I say, then, that belief is nothing but a more vivid, lively, forcible, firm, steady conception of an object, than what the imagination alone is ever able to attain" (Hume, An Enquiry Concerning Human Understanding, 49).

either true or false. In addition, since they are not products of reason, they are non-rational attitudes. Thus, we cannot say that Robert's statement is rational or irrational: "Not only are desires not assessable in terms of truth and falsehood, they are not subject to any sort of rational criticism at all" (Smith ibid., p. 8).[9]

Furthermore, another way of locating the difference between desires and beliefs is through the *directions of fit*. In *Intention*, G.E.M. Anscombe introduces the idea of directions of fits with the metaphor of a *shopping list* and *detective's record.* She invites us to consider the following,

> … a man going round a town with a shopping list in hand. Now it is clear that the relation of this list to things he actually buys is one and the same whether his wife gave him the list or it is his own list; and that there is a different relation when a list is made by a detective following him about. If he made the list itself, it was an expression of an intention; if his wife gave it to him, it has the role of an order. What then is the identical relation to what happens, in the order and the intention, which is not shared by the record? (Anscombe 1957, p. 56)

There are two lists in the example above, namely, the buyer's shopping list and detective's record. Notice that they have different goals, in the case of the buyer's list, the man strives to get items in the store to fit the items in his list. Suppose his list says: 2 bottles of milk, 1kg of beef, a loaf of bread, and 5 cans of Egger Beer. On getting to the supermarket, his task is to buy the items in order to correspond to his list. Whereas, in the case of detective's order, the opposite holds, the task here is to make his order fit to the items purchased in the supermarket. In other words, it is asked: In what way does the man's list differ from the detective's order? Anscombe replies as follows,

> It is precisely this: if the list and the things that the man actually buys do not agree, and if this and this alone constitutes a mistake, then the mistake is not in the list

---

[9] There are, however, conditions under which desires can be criticized, for example, in cases where desires are based on false beliefs. Smith rightly argues that: "Since I can be rationally criticized for having the belief, as it is false. I can be rationally for having the desire it helps to produce. The proviso is thus fairly minor: desires are subject to rational criticism, but only insofar as they are based on beliefs that are subject to rational criticism. Desires that are not related in some such way to beliefs that can be rationally criticized are not subject to rational criticism at all" (Smith 2000, p. 401).

but in the man's performance...; whereas if the detective's record and what the man actually buys do not agree, then the mistake is in the record. (Anscombe 1957, p.56)

Anscombe's goal is to point out the different relations to the world. According to her, our relations to the world follow different attitudinal patterns. On this analysis then, we can characterize the man's effort to get the items purchased to conform to his list as a *world-to-mind* pattern; and the detective's task of making his order fit to the items purchased as the *mind-to-world* pattern. However, what do these relations have to do with beliefs and desires? Mark Platts describes how this idea applies to the mental states,

Beliefs aim at being true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit the world, not vice versa. Desires aim at realization, and their realization is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing in the desire, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa. (Platts 1979, pp. 256 -257)

On the one hand, beliefs, which are products of reason strives to match with the world; and they are supposed to fit to the world. That is, they capture things as they are in the world. The propositional character of belief further fleshes this fact out namely, a representation that aims at being true. If belief, however, fails to hit its truth-target, the subject holding it is forced to adjust it to fit the world, or abandon it if the belief radically deviates from being true. On the other hand, desires are different attitudes that follow the state of affairs of the desired things (Pigden 2007, p. 201). They do not aim at fitting to the world; rather, they pull this world to conform to them. Suppose I say, I want Dornbirner Bulldogs to win this year's Ice Hockey league. On this analysis, my aim is the realization of my desire, rather than whether my desire is true or false, good or bad.[10]

---

[10] Unlike beliefs that "tend to extinguish if the world does not correspond with the content of your belief" (Pigden 2007, p. 201), desires tend to persist until the world conforms to them. Smith writes that "the difference between beliefs and desires in terms of direction of fit can be seen to amount to a difference in the functional roles of belief and desires […] a belief that p tends to go out of existence in the presence of a perception with the content not p, whereas a desire that p tends to endure, disposing the subject in that

The psychological states in question fundamentally differ from each other. On the one hand, moral beliefs describe how things are in the world morally. On the other hand, desires bring about certain changes in the world. Therefore, we can say that

> (P1): Moral beliefs and desires are distinct and independent psychological states, and there is no connection between them.

Assumption (P1) implies the following: if we claim that moral judgements are not beliefs, then they must be some sorts of desires; and vice versa. The non-cognitivist holds the former view. According to him, moral judgement expresses attitudes or feelings instead of beliefs. Non-cognitivism embraces two main views, viz., the *semantic thesis* and *psychological thesis*. The former maintains that moral judgement is not in the business of describing moral facts, and as such, they are not true or false propositions – *non-descriptivism*. In the latter, it is argued that moral judgement does not express belief*;* hence, the mental state supporting moral judgement is not cognitive in nature. On this construal, therefore,

> (P2): Moral judgements are not propositions, which are true or false.
> (P3): Moral judgements are not beliefs, hence not cognitive-laden.

Given that moral judgements are (P2) and (P3), it naturally follows that moral judgements are embedded on some sort of conative states such as desires, emotions etc. However, what practical implication does (P2) and (P3) assumptions have on moral practice? Precisely, how is moral motivation possible within the landscape of non-cognitivism? Since moral judgements express attitudes or desires, which bring about changes in states of affairs, the non-cognitivist argues that they are capable of motivating us efficaciously. Stevenson strongly argues that:

---

state to bring it about that p" (Smith, 1994, p. 115). However, there are cases where desires are not realized, for example, I have always desire to play the guitar, but the world has not yet conformed to the content of my desire. But, the fact that my desire is not yet realized, at least until this moment, does not mean that it is a false or bad desire.

A person who recognizes that x to be "good" must *ipso facto* acquire a stronger tendency to act in its favour than he otherwise would have had. (Stevenson 1937, p. 16)

This explanation seems to capture the connection between moral judgement and motivation. The idea of motivational internalism is often thought to be compatible with non-cognitivism because the expression of an agent's emotions or attitude is considered as an indispensable part of moral considerations (Brink 1986, p. 27; McNaughton 1998, p. 136).[11] Moreover, non-cognitivist motivational internalism seems to fit into Hume's psychology, because according to Hume, moral judgements "influence actions and affections…. they cannot be derived from reason; and that because reason alone, as we have already proved, can never have any such influence" (THN 457/294).

1.5. Tension

Assuming Hume's constraint on belief is true; does it mean that belief-based moral judgement does not have any practical implication on morality? Taken on this level, the most moral judgement can tell us is how to objectively and correctly answer moral questions and problems,[12] but it cannot move us to act accordingly. However, why is belief incapable of motivating us? According to Hume, the rationalist located moral judgement within the confines of relations of ideas; and as such, the most reason can only tell us is how we can deductively arrive at moral conclusions or principles. Precisely, beliefs, the products of reason, can only tell us how to get what we desire. In other words,

---

[11] On the non-cognitivist analysis, since moral judgement essentially expresses emotions or attitudes and not belief, it is maintained that if an agent fully understands the meaning of the concept of approval or disapproval intimately tied to his moral judgement, then it is inconceivable for the *connection* between moral judgement and motivation to fail. Alternatively, cases of failure are due to the agent's inability of understanding the moral terms contained in his moral judgement.

[12] There seems to be further ground on which to question the practical implication of objective moral judgements. Non-cognitivism, among other things, strongly contends that moral facts on which such objective moral judgements are based do not exist. Assuming this claim sticks, both the theoretical as well as the practical aptness of objective moral judgements in question will be threatened.

it has only an *instrumental use*.[13] Nevertheless, it can neither generate desires nor cause action:

> Morals excite passions and produce or prevent actions. Reason of itself is utterly impotent in this particular. The rules of morality, therefore, are not the conclusions of our reason [and hence, presumably, neither true nor false, since reason is our belief-forming faculty]. (THN 457/294)

Recall Dave's case, his correct answer can only tell him that stealing is bad. However, on Hume's analysis, it is not capable of making him not take the money in the absence of a fitting desire. Nonetheless, does it mean that belief-based moral judgement cannot motivate us at all? I do not think so; and if this were true, then moral realism will be impractical, hence motivationally impotent. On the contrary, I contend that,

(P4) Moral beliefs motivate.

The concern here is, essentially, how to explain how motivation is possible under moral realism. In other words, if the moral realist rejects (P2) and (P3), then he has to justify motivation in terms of moral belief. There are two main options open: [1] the moral realist can hold that motivation necessarily follows from moral belief; in this case, he will be espousing *Internalism*. Generally, the view that:

(P5) If an agent believes that it is right to φ, she will be motivated to φ.

On this construal, moral belief is not only a genuine belief, but it is argued that it can motivate us as well. In this work, I argue that this position presents an incomplete picture of human psychology. It presents an implausible solution to the moral motivational

---

[13] It is standardly argued that reason does not constitute ends, rather, it serves only as an instrument or compass that guides us to our ends. "Human reasoning involves the determination of means to achieving ends…These ends are in no way fixed by reason operating non-instrumentally; i.e., what makes them our ends is something other than reason…. Hume's famous remark "Reason is … that slave of the passion" (T415) is a clear endorsement of this thesis" (Hampton 1995, pp. 57 – 58).

problem.[14] First, if moral belief alone motivates us necessarily, I argue that the realist motivational internalist saddles the burden of explaining away the role of desires in moral actions in particular, and human psychology in general. Second, if he accepts that desires, at least, have a place in human psychology (no matter how insignificant it is), then he has to explain how moral beliefs generate those desires or how moral beliefs encapsulate those desires.[15] Assuming Hume's picture of human psychology is correct (P1), the realist motivational internalist has the task of explaining either how motivation is caused by moral beliefs alone or how desires figure in moral action. Of course, he has to offer a justification that does not trivialize the connection or compromise his internalist thesis. Notice that this task entails combining (P4) and (P5) while rejecting (P2) and (P3).

Third, by claiming that moral judgements are genuine beliefs; but unlike non-moral beliefs, they motivate us on their own; the realist motivational internalist shoulders the burden of justifying why moral beliefs possess both representational contents and motivational force. In other words, he has to justify the source of the normativity of moral beliefs (Bromwich 2009, p. 344). Since accepting realist motivational internalist solution seems to place moral realism under the threats of not sufficiently accounting for the practical implications of moral judgement, I adopt the second option, namely [2] *Externalism.* I argue that moral judgements are beliefs, but they do not motivate on their own. Thus,

> (P6) The source of moral motivation is external to and independent of moral judgements.

---

[14] By insisting, that moral realism is compatible with internalism, the realist motivational internalist is making the non-cognitivist attack on moral realism a real threat. We shall not allow ourselves to be forced into explaining the connection between moral considerations and motivation on a ground that is neither compatible nor derivable from moral realism. We shall not succumb to the seductive arms of internalism. Brink captures the idea I have in mind when he writes: "But the implausibility of conjoining realism and internalism […] derives from internalism and not from moral realism. Morality is practical, as our moral beliefs and practices assume, and so we should expect to find important connections between morality and both motivation and reasons for action. But internalism is not the correct way to represent these connections". (Brink 1986, p. 27)

[15] Further, if he argues that the psychological states are not distinct and independent of each other, then he has to accept that they are complex states. Even on this level, he cannot dodge the justification burden, because the account of how these complex states bring about moral actions will have to be justified.

Externalism holds that "the motivational force and rationality of moral considerations depends on factors external to the moral considerations themselves" (Brink 1989, p. 42). I defend motivational externalism against realist motivational internalism. I argue that by making moral motivation analytically dependent on moral judgement, the realist internalist removes the moral person from the landscape of moral agency. Precisely, the conceptual intuition on which realist internalism rests is incomplete, hence implausible, because it saddles the burden of explaining away, for example, the crucial roles of agent's desires, self, traits and identity in motivation. Further, it fails also to encapsulate the place of moral education in the agent's motivational profile and moral actions. Therefore, the objective of my dissertation is to fill this gap.

This work is divided into two major parts. In Part One, I consider the issues pertaining to the theoretical feature of morality. Chapter 2 focuses on noncognitivism. Contrary to the claim that morality is not about anything in particular, I argue that moral thought and talk are about *something* which is assessable in terms of truth and falsehood. I begin by arguing against moral sceptic claim – moral judgements are not factual. I argue that moral judgements are value-fact stating judgements, and that despite its specificity of normativity, morality is factual. It is factual in a different sense; and the moral domain has the epistemological resources for describing the realities of its domain (2.1). I then turn to moral expressivism. Precisely, on the one hand, I argue that the emotivist conative-based morality and its denial of truth-apt make morality objectionably dependent on agent's attitudes and fail to explain the semantic and logical features of moral discourse. On the other hand, the *internal reading technique* of hybrid expressivism does not give us a plausible picture of moral objectivity (2.2). Given that expressivism does not sufficiently capture our straightforward intuition about the theoretical feature of morality, I adopt a robust metaphysical stance, namely moral realism. In chapter 3, I, first characterise the defining features of moral realism to map off cognitivist and realist-like positions that do not entail moral realism (3.1 – 3.3). Second, I defend a nonreductive naturalist version of moral realism relevant for the purposes of this project (3.4).

Part Two considers the practical feature of morality. I attempt to achieve two main objectives: First, to argue against realist motivational internalism; and second, to expound

and defend externalism. In chapter 4, I argue that the phenomenon of moral indifference threatens the *motivational efficacy claim* of robust internalism. However, given that the internalism-externalism debate is construed in a number of various ways, I walk through these views (4.1). This stage-setting exercise enables us to identify the underlying features of motivational internalism; and thus, helps us to characterise the versions of realist motivational internalism relevant for the purposes of this work. I argue that motivational internalism based on the first-person judgement has an advantage over the types that tied motivation to moral properties or moral obligations (4.2). I discuss additional elements – Internality and Conceptuality – of motivational internalism (4.3). Having outlined the core assumptions of robust motivational internalism (4.4), I argue that the robust construal of realist motivational internalist solution is threatened by the phenomenon of moral indifference (4.5). Finally, I explain why moral indifference is possible and why it is a threat to realist motivational internalism (4.6).

In chapter 5, I turn to conditionalized realist motivational internalism. I discuss the assumptions of conceptual rationalism (5.1 – 5.2); and I show that Smith's rationalist defence of motivational internalism is false (5.3 – 5.4). Further, I consider Smith's internalist argument from practicality requirement, and I argue that his fetish argument against externalism fails (5.5). I conclude that conditionalized realist motivational internalism does not sufficiently establish the *necessary* connection between moral judgement and motivation (5.6). Furthermore, in chapter 6, I consider four main objections against externalism, namely objection from moral intuition (5.1); objection from Action-guiding and motive-giving (5.2); objection from off-colour moral term usage (5.3) and objection from amoralism (5.4). I attempt to rebut them accordingly. In chapter 7, I develop a realist externalist solution to moral motivation. The core claim here is that moral motives for acting on moral judgements are based on independent desires and the desire of being faithful to the moral-self is one of them. I begin by presenting the externalist framework for explaining moral motivation (7.1). Since, my version of motivational externalism is based on the idea of a moral agent as a moral person; I elaborate the distinctive ways in which morality is crucially significant to the human person, self and identity – moral identity (7.2). Further, I explore the two main

understandings of moral identity: Kantian-Kohlbergian and Blasian Perspectives (7.3). Drawing on Frankfurt's "Second-Order Volitions" and Blasi's "Moral-Self", I propose an externalist account of a widespread connection between moral judgement and motivation. On this view, motivation is based on moral identity. I argue that *moral care*, *moral-self* and *moral desires* are essential to moral motivation. Moral motivation does not stem directly from moral judgements. Rather, it is from an independent desire, namely, a moral desire for the realization of the moral values we care about, hence, desire to be faithful to one's moral-self (7.4). Finally, I argue that externalism is not condemned to explaining the motivation of virtuous persons in terms of desire *de dicto* as Smith alleges. In fact, it is possible, on the externalist construal, to account for a reliable connection between moral judgement and motivation in terms of *de re* desires (7.5).

PART ONE

Moral Objectivity

Chapter 2

Noncognitivism

If moral questions are not *about* anything, and we are not talking about anything when we discuss them – any special subject matter of morality – then what *are* we doing?

Schroeder, Noncognitivism in Ethics, 2010

2.0 Introduction

In the previous chapter we characterised the theoretical feature of morality essentially from the moral realist perspective. It is argued that moral reality underlies our moral thought and moral language. Precisely, we encountered questions about whether or not there are determinately correct answers to moral questions; and whether or not such answers are dependent on agent's beliefs, conditions, and circumstances. We shall return to these metaphysical concerns of morality in the next chapter; and in addition, we shall show the kind of *thing* moral reality *is*. But before then, there are other sets of core questions about morality. First, it is asked, for example, whether or not we can *know* the answers to moral questions. Recall Dave's case in chapter one, it is argued that their talk and disagreement about *wrongness* of taking the money from the dormant account is determined by some sorts of moral reality. Thus, moral realism claims that moral questions about rightness and wrongness are sustained by moral reality. Nonetheless, can we be so sure of knowing moral wrongness or rightness? For example, *can* we or *how do* we know the wrongness of, say, stealing? Maybe we might find the question about the wrongness of stealing to be too obvious. What about issues such as assisted suicide or gay/lesbian relations? A survey from Gallup's annual Values and Beliefs conducted in 2010 shows that Americans are sharply divided on these issues. Can we, for sure, *know* whether or not assisted suicide is morally wrong? For some people, answers to such questions are evident, because of, for example, their belief in some sorts of moral reality

or convictions; but for others they are not. These sorts of concerns are epistemological questions about morality. We shall return to them shortly. Second, there is another related concern, namely, about our moral thought and moral language. Assuming the argument for the wrongness of assisted suicide sticks, it is asked, then, do we think and talk about *something*? In other words, it seems that we disagree not only about moral questions, but also, whether or not our thoughts and talks about morality have any meaning at all.

These sets of concerns are at the very heart of metaethical debate. From our discussion in the previous chapter, we can anticipate the moral realist response to these questions, given their claims about correct answers which are sustained by moral reality. For moral realism therefore, morality is essentially about *something*; and when we think and talk about it – rightness or wrongness – we are thinking and talking about *moral facts*. Consequently, we think and talk about them because we can *know* them or at least, because their knowledge is possible. In addition, the idea of *correctness* presupposes that moral questions are about something, which can be true or false. Schroeder fittingly argues that:

> If moral questions are not, really, about anything, then there is no real question to be answered as to what they are about. Likewise, if moral questions are not really about anything, then there is no real question to be answered as to how we find out about it. And similarly, there would be no real questions as to how we manage to talk about it or think about it. (Schroeder, 2010, p.8)

Notice that if moral questions are about *something*, then we can, at least, attempt to describe the kind of thing it *is*. Recall the example of nicotine; we can list its properties such as being a hygroscopic, colourless oily alkaloid with 162.231556g/mol. But, can we describe the wrongness of, say, assisted suicide? The force of accepting the claim that morality is about something, among other things, implies that we can know the properties of wrongness or rightness; and at least, express them in our moral talks. In other words, the authority of such moral presuppositions seems to compel us into believing that moral statements can be true or false, even when they are not obvious to us. Thus, one of the strategies for evading this metaethical challenge is to deny that morality is about something; and this is the strategy of antirealism in general, and noncognitivism in

particular. From the foregoing, we can roughly depict moral noncognitivism as a form of antirealism, which maintains that morality is not about anything which our thoughts and talks refer to, or reflect. Thus, when we engage in morality, our moral judgements are not in the business of describing or reporting something. In this chapter, I focus on a species of noncognitivism, namely *moral expressivism*. I argue that it makes morality objectionably incoherent, because it fails to capture our intuition about the theoretical feature of morality. Consequently, it does not give us sufficient reason to abandon the claim that morality is about something, which we can think and talk about. However, before then, I will consider the epistemological concern; specifically, I argue that moral knowledge is possible. Accordingly, morality is about something, which we can, at least, know.

2.1. Moral Knowledge and Scepticism

The idea of an external world typically includes the possibility of knowledge; and the objectivity claim is tied to knowledge, because the objectivity claim of any domain makes sense only if the knowledge of such a world is possible. Generally, when we talk of knowledge, it is normally related to *something* – a particular thing, a class of things, or certain domains. When we say we know, for instance, a particular thing, such a claim presupposes its existence as well as the possibility of, at least, a being that knows it (and is capable of knowing it). Consider, for example, Peter saying the following: *I know that p.* He is not merely claiming the knowledge of (or capable of knowing) *p*, he is as well positing, if only roughly, the existence of *p*. The latter claims, that is, the existence claim is likely to trigger off doubts, especially, when someone thinks that the being of *p* is not so obvious or when he sees the nature of *p* as being mysterious or strange. In such cases, he might argue that *p* does not exist; therefore, Peter's claim is not true. In other words, *he does not know p because such thing, as claimed, does not exist.* Notice that the existence of *p-thing* seems to point to the possibility of knowledge of *p* as well as truth conditions of *p-belief*. Suppose, we claim that there is *t*, (where *t* stands for a domain); and it is supported by standard ordinary discourse. Let us further, for the sake of

explanation, say that there is *t*-knowledge. Just as in the case of Peter, claiming that t-knowledge is possible is likely to generate some sort of scepticism. Let us call it *t*-scepticism.

More so, scepticism normally raises the 'how' questions. Alternatively put, claims such as 'I know that p' or 'there is t-knowledge' invite the 'how' questions. For the present purposes, a distinction is in order here. We shall, thus, separate what Quassim Cassam (2007) dubs the 'how' questions from the 'how-possible' questions. We are not interested in 'how', that is, the steps Peter passed before he came to know p. Using Schroeder's analogy (2010) of a drinking fountain (which we fully *know works* when we press the button and water comes out, although we might still be figuring out how it works), we argue that 'we do know' something about morality. Notice that this analogy points to two epistemological questions, namely: "do we know" and "how do we know". The 'do we know' question, when posited in any domain of inquiry, casts doubt on the possibility of knowledge of the domain in question. For example, one can doubt whether mathematical knowledge or aesthetic knowledge is possible after all. As I understand it here, I take 'do we know' question to be more fundamental, because the 'how do we know' question seems to presuppose that the knowledge of the subject-matter in question is, at least, possible. In other words, if I do not know that drinking fountain works in the first place, the question of '*how* it works' will not be urgently necessary. Thus, our interest is not on the processes or steps involved in knowing. Rather, we are interested in whether or not moral knowledge is possible at all.

That said, 'Is knowledge of *p* possible?' 'Is *t*-knowledge possible?' If it is not, why is it so? Cassam speaks of obstacles: for him, the 'how-possible questions are obstacle-dependent' (Cassam, 2007, p. 129). Suppose there is an obstacle hindering the possibility of the knowledge of *p* or *t*-knowledge respectively; our main question is: Is such an obstacle impossible to overcome? It goes without saying that the *domain-scepticism* comes in different flavours and degrees – strong, moderate, weak etc. But they seem to share a feature in common, namely a (part or complete) rejection of knowledge of a given domain of human inquiry. Our concern is on a group of sceptics who deny the possibility of moral knowledge. For them, there is no such thing as moral knowledge. The argument

of the moral sceptics generally runs as follows: There is no moral knowledge, because the 'purported' something, namely moral facts do not exist in the first place, let alone being objective.

In order to place our discussion in context, let us identify one of the main obstacles motivating this sort of moral scepticism. Generally, it proceeds by setting a demarcation between moral domain and other domains. Suffice it to say that one of its special points of reliance is the scientific domain, because the latter is so conceived as an exemplifier of fact-stating domains. The strategy is seen in A.J. Ayer's *Logical Positivist Argument*. To Ayer's mind, moral judgements are considered as meaningless because they are not analytic or empirically verifiable (although, as we shall see later, Ayer claims that moral statements have a different meaning, hence they are not entirely meaningless). Implicit in Ayer's position is the claim that moral judgements possess something, which other judgements lack; and the presence of these special features disqualifies moral judgements from playing in the league of 'factual' judgements. To illustrate, the moral sceptic strategy can be presented as follows:

(1)   Moral domain is about values

(2)   Domains (e.g. natural sciences) are about facts

(3)   Knowledge is about facts

(4)   Therefore, moral knowledge is not possible

Just in case we accept the claim in (1), does it mean that there is no possibility of knowledge within the value-domain? Specifically, are value-domains, especially the normative domain, devoid of knowledge? I don't think so: the normative domain is not just about value judgements. It is also about factual judgement; after all, what are values outside facts? We can describe normative judgements, precisely moral judgements, as value-fact stating judgements. However, it is good to note that the fact-stating character of moral judgements differs from other factual judgements. It is not to be construed to be same as physical facts (or identical and reducible to natural facts). If we were to conceive the factual character of moral judgements in that manner, then we would be misconceiving the nature of moral facts. Alternatively, denying the possibility of moral

knowledge, simply because its facts are not same as those of other domains, is also a misconception. Every domain has its specificities. Even within the natural scientific domains, chemistry is distinguishable from physics and their objects differ respectively. While it is true that moral facts are radically different from, say chemistry; it would not be right to disqualify moral knowledge as 'knowledge' simply because of its specificities. We cannot claim that the knowledge of a given domain is not possible, simply because it lacks certain specificities, where the features in question – facts, methods, presuppositions, etc. – are special features of another domain. Railton was right when he made the following observations:

> The terms 'fact' and 'factual' should not mislead us. Our familiar idea of a fact is something like a state of affairs that one might encounter directly in experience, such as the presence of a concrete object at a particular place and time. This idea does not seem to fit the moral case. However, it also does not seem to fit many other domains where we speak readily of facts: facts of arithmetic, facts about what is probable or possible, or facts about the meanings of words or the causes of cancer. The nature of moral facts, if there are such, will depend upon the sorts of claims morality makes. (Railton, 2006, p. 202)

It is not surprising that a reasonable amount of moral sceptical assumptions is based on this misconception. Notice that this misconception also drives assumptions (2) and (3). These assumptions set moral knowledge against the facts of other domains. The main supposition is that the moral domain lacks the kind of facts that other domains possess; and given this, we cannot be so sure that moral knowledge is possible.

For the time being, however, let us assume that there is, at least, a single true fact, say 'G', in the moral domain. We can show by parity of reasoning that just as this fact will not fit into any empirical domain; so also, the facts of other domains will not fit the moral domain. We cannot say that knowledge is impossible in some domain 'S', because it does not fit into our moral fact. In other words, we cannot deny a given domain its knowledge-claim simply because it does not fit into a purported single domain. I refer to this as the 'story of a single domain'. Notice that moral sceptics seem to apply single-

domain argument, whereby the moral domain is supposed to conform to the specificities of the purported factual domain. However, the main question to ask is not whether the moral domain is factual on the standards of a given 'purported' fact-stating domain. Rather, whether the moral domain has the epistemological capacity of describing the realities of its domain. Moreover, if we sustain our claim about the possibility of a true moral fact 'G', then we can rightly ask, 'Are the epistemological resources of moral realism capable of describing 'G''? Railton writes the following:

> Ordinary factual statements *describe* how things are, past, present, or future. Moral language employs directive and evaluative concepts, and moral statements make claims that *prescribe* or *command*, saying what should or shouldn't be done, or how things might be better or worse. (Railton, 2006, p. 201)

He further asks, although

> Moral language is therefore seen as *normative* or *action-guiding*. Could it nonetheless be factual? Could saying how things might be better or worse, or what moral obligations an individual has, also be a way of saying how things *are*? (Railton 2006, p. 201)

Despite its specificity of normativity, morality is factual. It is factual in a different sense; and its epistemological resources can describe moral facts. That is, it can say how things are morally. Moral practice shows that moral language not only asserts something about its subject, but also, it makes propositions, which are truth-apt. In other words, we can, as we argued in chapter one, discover moral correct answers through moral reasoning or deliberations, because there are facts supporting our moral language. So, it is not correct to say that:

(5) If moral knowledge *is* unlike other domains, it is not knowledge.

(6) Therefore, moral knowledge is not possible.

Notice that assumption (5) is built on *sameness* and not on *similarity*. The difference between moral domain and other domains is clear, just as the difference between an apple and a pear is clear. Nevertheless, it is not true to claim that pears are not fruits simply

because they are not in the same way as apples. Morality has its objects and facts, just as every other domain of human inquiry. While we do not expect morality to be the *same way* as other domains; however, we expect the moral domain to be *similar* (to other fact-stating domains) especially in describing its facts: This is the place of objectivity in morality. In other words, the objectivity of moral judgement lies in the fact that it describes its facts in a way similar to other fact-stating domains. Moral knowledge is possible if our moral belief does not cause moral values; and if it tersely captures these values or facts of morality without depending on feelings or projection of attitudes.

Furthermore, the moral sceptic argues from *certainty*. He is *certain* that there is no such thing as moral knowledge. At least, most of the moral sceptics believe that there is an external world, and this is seen in their reliance on the scientific domain (as a means of rejecting the possibility of moral knowledge). Based on this reliance, they hold it to be *certain* that their claim against the possibility of moral knowledge is *true*. For example, if we ask a sceptic of this class, 'Can you explain why there is no such thing as moral knowledge?' He (literally) goes like this 'We *know* it to be *true* that the so-called moral knowledge does not deal with facts but values'. Notice that this position begs the question, because the explanatory tools he employs are not from his 'purported' fact-stating domains either (we shall see this replicate in the emotivist claim below). In other words, by saying that there is no moral knowledge, he is positing a conceptual claim *about* knowledge – meta-claim about the knowledge of the impossibility of moral knowledge. That is, he is saying that he *knows* that moral knowledge is not possible. However, which *fact* supports his claim? Is his claim (empirically) factual? If not, we can deny his knowledge-claim on the ground of its non-conformity to his 'purported' fact-stating domain. Nevertheless, if he insists on sustaining his certainty stance, then, we have grounds to submit that moral knowledge is possible as well. He faces a dilemma here.

In a similar context, Bambrough (1979) notes that moral sceptics accept the fact that common sense supports scientific domains, but they argue that morality lack such a support. If the moral sceptic is to be taken seriously, he has to justify how his own talk is supported by standard ordinary discourse. If the moral sceptic, however, accepts the possibility of an external world, then we can prove that moral knowledge is supported not

only by common sense, but also, by ordinary language. In other words, if we are able to establish that there is an external world, "then we can show *by parity of reasoning*, […] that we have moral knowledge, that there are some propositions of morals which are certainly true, and which we know to be true" (Bambrough 1979, p. 15). For example, propositions such as *it is morally wrong to torture a child for sexual gratification*,[16] is not only morally true, but common-sense supports it as well. If Bambrough was right in submitting that no force of reasonable argument can disprove the truth of such propositions, then it seems not wrong to suppose that there are moral facts that support the strength of the propositions of morals in question.

Finally, while we might listen to Hume's advice [namely: "The only way, therefore, of converting an antagonist of this kind, is to leave him to himself. For, finding that nobody keeps up the controversy with him, it is probable he will, at last, of himself, from mere weariness, come over to the side of common sense and reason" (Enquiries, §133)], we still expect him to disprove the truth-condition and truth-maker of moral propositions (in our own here: *it is morally wrong to torture a child for sexual gratification*). In sum, he has to explain away why such knowledge is impossible.

2.2.Moral Expressivism

The cognitivist-noncognitivist debate about morality cuts across philosophy of language and philosophy of mind, however, for the purposes of this work, I will consider them as the theory of what we do when we engage in moral discourse, because thought and language are closely tied. Schroeder puts it as follows:

> But just as we can wonder how you manage to *find out* about this thing – that is, how you *know* what is right or wrong – we can also wonder how you manage to *talk* and *think* about it. The question of how you manage to talk about it is a question from the philosophy of language, and the question of how you manage

---

[16]Bambrough used a different example.

to think about it is a question in the philosophy of mind. These questions sound very similar, and in fact they are closely related. (Schroeder, 2010, p. 7)

We had characterised noncognitivism as the view that morality is not about anything which our thoughts and talks refer to, or reflect. Precisely, it claims that when we engage in moral thought or talk, we do not assert things which can be true or false. In essence, our moral judgements are not beliefs; therefore, they are not truth-apts. On this model, the statement as we saw above: *It is morally wrong to torture a child for sexual gratification*, is neither true nor false. For noncognitivism, truth-maker does not exist. This is one side of the noncognitivist story – its negative claim. On the other hand, noncognitivism tells us that when we make moral judgements we express *pro* and *con* attitudes. Thus, the statement: *It is morally wrong to torture a child for sexual gratification* expresses the speaker's disapproval of the action. Given its various manifestations, the positive claims of noncognitivism differ. Various noncognitivist positions amount to different postures to moral discourse; and they link it to different attitudes. As a consequence, their manifestations differ in various ways, thereby producing a diverse family of views. In what follow, I will focus on moral expressivism; first, because it is one of the initial representations of noncognitivist reactions against the claim that moral language behaves like ordinary language in describing facts. Second, expressivism, especially, in its recent characterization is ambitious, because it promises to reconcile expressivism and realism in ethics; thereby giving us grounds for avoiding the metaphysical burden associated with construing truth-makers of moral statements robustly. (Notice that on this construal, *it is morally wrong to torture a child for sexual gratification* can be true or false – it is the truth-maker that it denies but not the truth-value of the statement). However, I argue that hybrid expressivism fails in its project. Although it accommodates truth-value into its expressivist claim, it does not tell us a plausible story about moral objectivity. But before then, I will examine expressivism in its traditional form – *emotivism*. I argue that it fails to capture the semantic and logical features of moral discourse, thus, not fit for explaining the theoretical feature of morality.

Emotivism

Generally, expressivism holds two main positive claims:

E1: Moral judgements express conative attitudes.

E2: The psychological states expressed in moral judgement are conative-laden.

[E1] and [E2] are the *semantic* and *psychological* claims of expressivism. However, it is a bit vague depicting expressivism as a view which holds that moral language expresses conative states of mind. What the expressivist posture entails is subtle, because it claims that the noncognitive states expressed in moral judgments are not asserted when agents utter approvals or disapprovals toward something. The alternative view would imply that moral judgements express agent's attitude toward 'X' as well as report his states of mind when he is approving or disapproving 'X'. For example, in uttering that '*assisted suicide is wrong'*, he is reporting that he has a negative attitude toward assisted suicide. This characterization seems to fit some versions of moral subjectivism, but not expressivism (at least in its traditional construal) as it would force it to accept what it wants to avoid, namely truth-aptness. Notice that if moral judgements both express an agent's attitudes and report his mental states, then the former can be *false* just in case the latter is not appropriately reported, and vice versa. For the expressivists, however, when we use moral terms like wrong or right in moral utterances, we are essentially expressing noncognitive attitudes (e.g. attitudes, intentions, emotions, wishes etc.) Specifically, we can impute the following claim to expressivism, namely that it is not a metaethical theory *about* states of mind, but *about* what we do with language when we engage in moral thought and talks; and this idea is clearly expressed in emotivism, a traditional construal of moral expressivism.

Ayer's Emotivism

Essentially, A. J. Ayer argues that when we use moral terms in utterances, they do not have literal significance as in the case of descriptive statements. He puts it thus:

> The presence of an ethical symbol in a proposition adds nothing to its factual content. Thus if I say to someone, 'You acted wrongly in stealing that money,' I am not stating anything more than if I had simply said, 'You stole that money.' In adding that this action is wrong I am not making any further statement about it. I

am simply evincing my moral disapproval of it. It is as if I had said, 'You stole that money,' in a peculiar tone of horror, or written it with the addition of some special exclamation marks. The tone, or the exclamation marks, adds nothing to the literal meaning of the sentence. It merely serves to show that the expression of it is attended by certain feelings in the speaker.

If now I generalise my previous statement and say, 'Stealing money is wrong,' I produce a sentence which has no factual meaning – that is, expresses no proposition which can be true or false. It is as if I had written 'Stealing money!!' – where the shape and thickness of the exclamation marks show, by a suitable convention, that a special sort of moral disapproval is the feeling which is being expressed. It is clear that there is nothing said here which can be true or false. (Ayer, 1952, p. 107)

From the foregoing, we can tease out what (Ayer's) emotivism amounts to: It does not quarrel with the structure of syntax of moral statements. While, for example, 'stealing money is wrong' might appear as though it has descriptive function, what it actually does is radically different. It expresses emotions and nothing more. Ayer advises us not to be confused by the syntax of moral statements, because it has no factual significance. Hence, on the emotivist construal, Lotta from Dave's case is just expressing her disapproval. She used the utterance: 'stealing money from the bank is wrong', to evince her moral disapproval of it; and such expressed attitude does not admit truth or falsity. Given this, Ayer claims that moral statements have *emotive significance* and not *literal significance*.

As a logical positivist, Ayer endorses the verification principle, the view that sentences have literal significance, hence descriptive function, just in case they are analytic or empirically verifiable.

The criterion which we use to test the genuineness of apparent statements of fact is the criterion of verifiability. We say that a sentence is factually significant to any given person if, and only if, he knows how to verify the proposition which it purports to express – that is, if he knows what observations would lead him, under certain conditions, to accept the proposition as being true, or reject it as being false. If, on the other hand, the putative proposition is of such a character that the

35

assumption of its truth or falsehood, is consistent with any assumption whatsoever concerning the nature of his future experience, then, as far as he is concerned, it is, if not a tautology, a mere pseudo-proposition. (Ayer, 1952, p. 35)

Based on this, propositions from domains such as metaphysics or ethics do not fulfil the criterion of verifiability. However, it is a bit strange that Ayer discarded metaphysical propositions as mere nonsense, but assigned emotive meaning to ethical propositions. The logical consequence of failing the criterion of verifiability would have been to eliminate metaphysical as well as ethical propositions. However, "it is less easy to see how Ayer can avoid this unwelcome implication: if ethical judgements are not literally significant, why doesn't ethics go the way of metaphysical nonsense?" (Miller 1998, 104)

Ayer's position, among other things, attracts the *why* as well as *how* questions. To the former, why did he assign emotive meaning to ethical statements? Although, he rejected the naturalist claim, namely, 'ethical statements do not describe how the world is', his motivation for assigning emotive meaning to ethical statements is unclear, given his epistemology. On his epistemological model, things we can know are the ones we can either analytically or empirically verify. However, assigning meaning to ethical statements, which is outside the realm of verifiability, is dubious. As we shall see in a moment, he adopts what Ronald Dworkin (1996) calls 'Archimedean' posture about morality. That is, standing outside the moral domain – namely that ethical statement is not the sort of thing we can verify – and at same time ascribing (emotive) meaning to it. Further, the latter is also linked to his epistemology: Does the verification principle support his claim that ethical statements do not have literal significance? Precisely, how does he know that 'ethical statements have emotive meaning'? At least, we expect him to be faithful to his statement, namely: "In putting forward the theory I was concerned with maintaining the general consistency of my position" (Ayer, 1952, p. 20). But it seems his claim about emotive meaning of ethical statements does not fulfil the criterion of verifiability.

Nevertheless, one might argue that the verification principle applies only to statements of facts. Accordingly, sentences like ethical statements can acquire their

meaning through other principles or procedures. After all, Ayer later relaxed his claim about meaningful sentences thus:

> In putting forward the principle of verification as a criterion of meaning, I do not overlook the fact that the word 'meaning' is commonly used in a variety of senses, and I do not wish to deny that in some of these senses a statement may properly be said to be meaningful even though it is neither analytic nor empirically verifiable. (Ayer, 1952, p. 15)

Suppose ethical statements are meaningful in this regard, Ayer is silent about by what principle(s) an ethical statement acquires its meaning. Ayer shoulders the burden of explaining away why ethical statement, even though it does not fulfil the criterion of verification, has emotive meaning at all; and how to test as well as identify sentences that are meaningful outside the realm of analytic and empirical verifiability.

Furthermore, assuming he insists that his claim about emotive significance of ethical statements holds, then he will be making a claim about *something* – about ethical statements – which can be true or false. To illustrate:

(1) Ethical statements are emotively meaningful.

   Thus,

(2) 'Stealing is wrong' expresses the speaker's disapproval of it.

On the contrary, suppose we claim that:

(3) Ethical statements are not emotively meaningful.

   Thus,

(4) 'Stealing is wrong' describes something in the world.

If Ayer is serious about his claims in (1) and (2) he will reject (3) and (4) as false. At least, his meta-claim about ethical statements is accessible in terms of truth and falsity. I doubt very much whether, Ayer or the expressivists in general, will claim that (1) and (2) are expression of emotions or feelings. In other words, he is asserting or denying *something* about the subject; and drawing on Schroeder's worry, we can ask the emotivist: 'Assuming our talk about morality is not about *anything* in particular, what then, are we doing when we talk and make claims about any subject matter of morality?' It is odd, if not practically useless, to engage in an enterprise that is not about anything at all. In other words, if morality (e.g. moral questions, moral claims, moral discourse, etc.) is not about *something,* which can, at least be true or false in an unqualified sense, then the debate between the cognitivists and noncognitivists will not get off the ground in the first place.

On a similar note, Dworkin argues, as we saw above, that although the emotivist denies that moral claims are genuine propositions – they are not about facts, but he still makes claim *about* and *in* the domain which he rejects. It is strange Dworkin argues "to stand outside a whole body of belief, and to judge it as a whole from premises or attitudes that owe nothing to it" (Dworkin, 1996, p.88). Precisely, the emotivist argues that moral wrongness or rightness is not about anything in particular, but he goes on to make claims *about* moral statements. For example, Ayer claims that (2), but denies (4). The emotivist accuses his opponent of making substantive claim about the wrongness of stealing. However, Dworkin rightly argues that the emotivist, by rejecting (4) is also making a substantive ethical claim. Even when the emotivist argues that, 'stealing is wrong' expresses the speaker's disapproval and nothing more, he is making a claim in the moral domain. Notice that the emotivist strategy includes, first, claims about moral judgement – stealing is wrong; and second, denial of metaethical claims of moral judgement – stealing is wrong is not accessible in terms of truth or falsity. Dworkin strongly contends that substantive ethical claims do not come apart from metaethical claims, because the latter is an expansion of, and continuous with the former. However, if the emotivist wants to retain his initial position, he has to give up describing his position as a metaethical theory, and settle with the normative ethical claim. In other words, he cannot maintain his Archimedean posture, while making claims *about* and *in* the moral domain. "Any

successful – really, any intelligible – argument that evaluative propositions are neither true nor false must be internal to the evaluative domain rather than archimedean about it" (Dworkin, 1996, p. 89).

Nonetheless, Ayer's immediate concern, as we saw above, is that moral terms such as *wrong* or *right* appearing in statements are mere ejaculations, and add nothing to the factual meaning of the statement. Rather, they purely express the speaker's disapproval or approval; and not in the business of describing things. Notice that the speaker's attitudinal stances are not determined by anything, but solely depend on the speaker possessing the desires or feelings *to do* or *shun* certain kind of acts as wrong or right. This posture commits the emotivist to making moral rightness or wrongness depend on the agents. We can illustrate this sort of moral mind-dependency as follows:

E3: In judging that (X) is morally wrong, an agent (A) expresses an emotional stance toward (X).

The wrongness of (X) depends on (A) possessing relevant attitudes toward (X). It is not the independent value of (X) that makes judgement morally wrong, but the agent's relevant attitudes toward (X). Let us replace (X) with "torture is morally wrong" and (A) with Anna. By saying 'torture is morally wrong', Anna is expressing her emotions, feeling or attitude of disapproval toward torture. In Ayer's words, Anna is saying: **Boo to torture**! Let us further substantiate 'torture is morally wrong' by saying that 'torturing (illegal) immigrants is morally wrong'. Suppose Anna witnessed a group of young people torturing two immigrants. By judging the action as morally impermissible, Anna is only expressing her emotional attitude, hence, her disapproval toward the act.

If Anna were to lack the relevant attitude she would not only be neutral in the face of the act but also, torture will neither be impermissible nor permissible in itself because the moral value of things depends on her judgement. This attitude raises the problem of explaining and communicating morality objectively. In other words, emotivism reduces morality to a subjective attitude. Given this, moral judgment would only follow the *world-to-mind* direction of fit. However, it would amount to relegating the theoretical feature of morality to the background. If the directions of fit are supported by ordinary linguistic and thought practice (as we saw in the previous chapter), then the emotivist saddles the

burden of explaining how his construal of moral judgement, which is conative in nature, is supposed to satisfy the objectivity of morality.

Furthermore, if there are cases (either actually or hypothetically) where agents make moral judgements without expressing emotions or feelings, then the claim that moral judgement expresses or evokes feelings or emotions will be false. Suppose Anna was recently diagnosed with some sort of emotional condition that blacks her out momentarily. However, nothing changed in her practice of forming moral judgements. Upon further investigation, it is also discovered that her condition does not affect her cognitive capabilities. If Anna still judges that something is morally wrong or right, even during her emotional blackout sessions, then the psychological state underlying her judgement is not purely conative; hence her moral judgement would not just express emotions. Assuming the emotivist insists that other psychological states do not figure in moral judgement, he still has to make sense of directions of fit, thus, the functions of the psychological states – beliefs and desires – required in moral practice. Apart from the above concern, we shall see in Frege-Geach Problem that emotivism faces the challenge of explaining the semantic and logical features of moral discourse. Thus, it fails to tell us a plausible story about the theoretical feature of morality. But before we return to this concern, let us consider Charles Stevenson's version of emotivism.

Stevenson's Emotivism

Unlike Ayer, Stevenson's emotivism was not borne out of strict verficationist concern; but like Ayer, he was concerned with the meaning of moral statements. Specifically, his motivation was due to his dissatisfaction with some accounts of moral subjectivism, namely the interest theories.

> Traditional interest theories hold that ethical statements are *descriptive* of the existing states and interests – that they simply *give information* about interests. … It is this emphasis on description, on information, which leads to their incomplete relevance. Doubtless there is always *some* element of description in ethical judgments, but this is by no means all. Their major use is not to indicate facts but to *create an influence*. Instead of merely describing people's interests they *change*

or *intensify* them. They *recommend* an interest in an object, rather than state that the interest already exists. … The difference between the traditional interest theories and my view is like the difference between describing a desert and irrigating it. (Stevenson, 1937, p. 16)

From the above, Stevenson thinks that the descriptive emphasis on these theories inadequately captures the main function of ethical language. For him, construing ethical language as a mere fact-stating tool is a deep misconception, because it does more than that. We can already see that Stevenson departs not only from the traditional interest theorists, but also from Ayer with regard to moral linguistic functions. Unlike Ayer, at least, Stevenson accepts that ethical statements can *tell* us *something about* persons, situations or actions. Do note that the descriptive function of ethical statements is not epileptic: 'Doubtless there is *always* some element of description…' (Emphasis added). Nevertheless, Stevenson does not understand the descriptive aspect of ethical statements essentially as a truth-apt, given emotivist commitment as we shall in a moment. Further, he accepts that moral terms such as wrong, right, good, bad etc., figuring in utterances are not literally meaningless. They add meaning to utterances, nevertheless, ethical language fundamentally serves to express a speaker's *attitude* as well as *create influence* on the hearer. In other words, moral terms have emotive meaning(s).

> The emotive meaning of a word is a tendency of a word, arising through the history of its usage, to produce…affective responses in people. It is the immediate aura of feeling which hovers about a word. (Stevenson, 1937, p. 23)

Stevenson's main concern is not necessarily to repudiate interest theories in their entirety, but he thinks their account of 'good' is insufficient. According to him, claiming that 'good' means 'desired by me' (Hobbes) or 'desired by many' (Hume)[17] is in part relevant; but they cannot offer plausible explanations of: (1) moral disagreement, (2) moral motivation and (3) autonomy of ethics. For example, we disagree about 'what is good', but Hobbes' account makes moral disagreement impossible. "For consider the following

---

[17] Here both Ayer and Stevenson agree with Moore's charge against naturalism, the view that any attempt to define 'good' in terms of any natural property leads to naturalistic fallacy.

argument: 'This is good.' 'That isn't so; it's not good.' As translated by Hobbes, this becomes: 'I desire this.' 'That isn't so, I don't.' The speakers are not contradicting one another, and think they are, only because of an elementary confusion in the use of pronoun" (Stevenson, 1937, 16). More so, if 'good' serves only to describe or give information as the traditional theorists claim, Stevenson argues that the descriptive aspect of 'good' cannot move the speaker or hearer to act accordingly. Description lacks dynamism, because "a person who recognizes X to be 'good' must ipso facto acquire a stronger tendency to act in its favor than he otherwise would have" (Stevenson, 1937, 16). In addition, Stevenson argues that the traditional interest theorists misconceived ethics as a scientific discipline. Given that there is more to ethics than just reporting, "the 'goodness' of anything must not be verifiable solely by the use of scientific method" (Stevenson, 1937, 16).

Stevenson's main interest is, if only roughly, to irrigate the desert of the traditional interest theories. He claims to offer a new version of interest theory that makes up for the deficiencies of the traditional interest theories.

> I believe that the three requirements, given above, are perfectly sensible; that there is some one sense of 'good' which satisfies all three requirements; and that no traditional interest theory satisfy them. But this does not imply that 'good' must be explained in terms of a Platonic Idea or of a Categorical Imperative, or of a unique, unanalyzable property. On the contrary, the three requirements can be met by a kind of interest theory. But we must give up a presupposition which all the traditional interest theories have made. (Stevenson, 1937, 18)

Stevenson's solution is this: he thinks that the traditional interest theorists omitted the fundamental function of ethical language, namely the emotive function. Although, ethical statements have some elements of description, they are fundamentally used to express attitudes or feelings. For example, by saying that 'Abortion is wrong' the speaker is expressing his disapproval of it. In addition, moral judgements create influence in others. Hence, by saying that 'Abortion is wrong' the speaker also wants others to 'feel' so as well. (Stevenson, 81) Nonetheless, left on this level, emotivism appears to be committed to a deep subjectivist stance: Moral judgements might be understood as solely expressing

the speaker's attitude. In *Ethics and Language*, Stevenson addressed this concern as follows:

> 'This is good' has the meaning of 'This has qualities or relations X, Y, Z …,'except that 'good' has as well a laudatory emotive meaning which permits it to express the speaker's approval, and tends to evoke the approval of the hearer. (Stevenson, 1944, p. 207)

On this analysis, 'good' has the qualities or relations of X, Y, Z (hence, descriptive element), but principally it has emotive function. Thus, by saying that, 'Abortion is wrong' the speaker is making a claim about *something* – abortion which he disapproves and wants others to disapprove as well. For example, suppose 'abortion' as the qualities of 'A' and 'B', the speaker by saying that 'Abortion is wrong' is picking out those qualities in his disapproval of terminating a pregnancy. The speaker's attitude of disapproval, as I understand Stevenson, seems to follow from 'A' and 'B' qualities. However, I am not sure whether an emotivist can consistently make such a claim. Understandably, such a claim allows him to avoid the mind-dependency of subjectivism, but it is unclear whether attitudes being consequential on the qualities or relations of 'good' presuppose, or at least include, some sort of determinacy. In other words, it is asked whether there are distinct qualities or relations, actions, or situations as well as distinct corresponding attitudes one must possess before we can make moral judgements. Stevenson seems to be saying that in making a claim about abortion, the speaker is not making a claim about himself or his attitude, but about (the qualities of) abortion which he disapproves. This raises some questions, first, we might expect him to say something about the nature of such qualities or relations, whether they are moral or non-moral. Second, we might require him to explain how a speaker can express attitude about something (which has qualities or relations) without referring to it. We shall return to the latter concern because Blackburn seems to be making a similar claim. But before then, among other things, Stevenson argues that emotivism can offer a plausible account of moral disagreement. However, I argue that his argument does not fit our ordinary understanding of disagreement.

Moral Disagreement

Recall Hobbes' definition of 'good' as 'I desire this'. This definition can be extended to 'my community likes this'. Notice that the definition includes the speaker; it is speaker-inclusive. We can contrast this with the definition, 'good' means 'most people like'. Here, we assume that the speaker refers to what a group of people likes, which he is presumably not part of, hence non-speaker-inclusive. As I understand him, Stevenson's charge is mainly against the speaker-inclusive cases such as Hobbes' "I like this." "This isn't so, I don't like it" do not allow for a sensible disagreement. If 'good' becomes and depends solely on what the speaker desires (that is, when he speaks about his own attitude or himself), then disagreement would readily disappear. Consider,

(1)

A: I like X

B: I don't like X

Both Ayer and Stevenson argue that disagreement is not possible in (1), because it merely reports the speakers' respective attitudes towards X. They are making claims about themselves and their attitudes respectively. While Ayer thinks that genuine moral disagreement is impossible, Stevenson argues that: "In the first place, we must be able sensibly to disagree about whether something is "good" (Stevenson, 1937, 16). Although, Stevenson thinks that moral disagreement is possible, there is a sense in which he denies the possibility of moral disagreement, because of his non-cognitivist commitment. If moral statements are not truth-apt, we cannot disagree about issues which are presumably supported by moral facts, because there are no such things. Stevenson's goal is to offer an alternative version of moral disagreement.

First, he begins by separating what he calls 'disagreement in belief' from 'disagreement in attitude'. When people disagree in the former case, they hold incompatible positions in what they *believe*. However, when people disagree in the latter, they express *for-attitudes* and *against-attitudes* about something. Consider,

(2)

Jan: I want to eat hamburger

Kelvin: I don't want to eat hamburger, rather Kebab.

In (2) the cause of disagreement, unlike opposition in belief, is the speakers' preference. Stevenson writes that "the men have divergent preferences and each is trying to redirect the preference of the other – though normally, of course, each is willing to revise his own preference in the light of what the other may say" (Stevenson, 1998, p. 44). Given that Stevenson argued, as we saw above, that ethical statement is essentially emotive in its function; his main concern is to show that moral disagreement consists essentially in attitude and not in belief.

> When people argue about what is good, do they disagree in belief, or do they disagree in attitude? … It must be readily granted that ethical arguments usually involve disagreement in belief; but they also involve disagreement in attitude. And the conspicuous role of disagreement in attitude is what we usually take, whether we realize it or not, as the distinguishing feature of ethical arguments. (Stevenson, 1998, p. 45)

He thinks we are mistaken in construing moral disagreement in terms of beliefs which can be true or false, because although moral statements may appear as assertions, but they are dynamic. Thus, when we disagree morally, we are just disagreeing in attitudes about something. However, this idea does not seem to shift our straightforward intuition about ordinary disagreement in general and moral disagreement in particular. In both cases, when we disagree, we assume certain epistemic postures toward moral statements. If only unqualifiedly, we take (moral) statements to have truth-value. We know, for example, that we can be mistaken about the properties of nicotine or unsure whether assisted suicide is wrong. To resolve such issues, we need moral statements to have truth-value. In other words, we consider people holding contradictory statements to be genuinely disagreeing both in moral and ordinary senses.

In addition, when two people disagree on a moral issue, we expect them to be arguing about the same object. This experience is, as well, prevalent in the ordinary discourse. Consider,

(3)

C: Snow is white.

The statement reports the property (whiteness) of an object (snow). Disagreement, in turn, sets in when someone denies C's statement,

D: Snow is *not* white.

(C) and (D) refer to the same object. They are asserting something about the property of snow, namely its whiteness. Both statements cannot be right at the same – either C or D must be wrong. C disagrees with D because he believes that D's statement does not report the object properly. We disagree because we believe someone must be mistaken; and the disagreement will persist just if he refuses to give up his position. This issue points to a further feature, that is, we expect C and D to stick to the same object as long as the disagreement lasts. C and D will be talking past each other, should they shift to another object or fail to properly define the object under consideration.

However, it seems that Stevenson's account of moral disagreement in attitude fails to meet these requirements. When people express a pro-attitude 'X', we assume that they are 'X-ing'. At least, by default they have a similar or share in the 'X-ing' experience. For example, we expect people 'loving' to be expressing a distinct attitude of 'love' by default. Suppose two people express their love for a particular object 'P',

(4)
Jan: I love 'P'
Kelvin: I love 'P'

They have the same object 'P', but it is not clear whether or not we can say that they are expressing the same attitude towards 'P', at least, given the subjective features of

possessing attitudes. By contrast, suppose D sees that he is mistaken and consequently, gives up his position.

D: Snow is white.

Just in case there is such a property (whiteness) out there (or we presume that such a thing exists); hence, C's statement is true, because it captures it tersely. We can be confident that D's statement in (4) not only refers to the same property, but also his statement has the same belief-content as C's. For example, when we assert, we are reporting something independent of what we think or feel. In the case of attitudes, although Jan and Kelvin seem to be directing their attitudes to 'P' – same object – subjectively they might be expressing different attitudes to 'P'. Given that attitudes, unlike beliefs, do not pick out something in the world as we saw in the directions of fit in the previous chapter, we cannot be so sure whether or not their respective attitudes capture 'P'. There is a psychology supporting both beliefs and attitudes respectively. We are not claiming that the psychology of belief does not admit subjective or qualitative difference in agents, say C and D. What makes the difference is, regardless of the differences, their beliefs aim at the same object outside of the speakers. Hence, the correctness of their statements does not depend on what they think or feel, but on picking out the object in question as properly as possible. Further, Suppose

(5)
Jan: I love 'P'
Kelvin: I don't love 'P'

On Stevenson's analysis, they are disagreeing about 'P'. Precisely, he argues that their disagreement is due to opposing attitudes. This position raises some worries: when we disagree, can we properly say we 'argue' about attitudes? Since attitudes do not have truth-value, how can we resolve the disagreement in (5), supposing there is one in the first place? Although, Stevenson offers a solution, however, as we shall see shortly, he

begs the question. Further, it is possible that in (5) Jan and Kelvin are expressing their emotions towards 'P' with subjectively different attitude-contents; and this seems to have an "Ought implies Can" implication. Morality should make claims on an agent insofar as she *can* do what it demands; otherwise it would be "pointless to place a person under a moral obligation if she is unable to fulfill it" (Kühler and van Ackeren, 2016, p.5).[18] However, it seems that the emotivist morality would lead to moral demands that (at least some) people cannot fulfil, especially in terms of feeling or attitudes. Consider,

(6)

Jan: X is good

Kelvin: X is bad

Using Hall's model (1947), we can translates this into,

Jan approves X; Kelvin, do approve X as well!

Kelvin disapproves X; Jan, do disapprove X as well!

This amounts to declaratives as well as imperatives – they both express their attitudes, while commanding each other do so well respectively. In addition to Hall's contention that there is no direct or implied contraction between them, hence no disagreement in (6), it is unclear whether Kelvin can do what Jan requires of him and vice versa. For example, by commanding Kelvin to X, Jan wants him to share the same attitude toward X. Notice that it seems impractical, if not impossible, for B to possess (or at least assume) A's attitude given the subjective features of individual's attitude. If we assume that we cannot simply develop our subjective attitudes (especially an emotionally laden one) at will, then it appears, for example, that Jan is demanding Kelvin to do what he is unable to fulfil. Even if it is possible by default, we cannot be so sure of its outcome, since attitudes do not terminate in objects which we can be described, hence be true or false.

Furthermore, Stevenson denied that moral judgements do not report the speaker's attitude, so there is no means to knowing whether speakers' attitudes are right or wrong.

---

[18] For more discussion on "Ought implies Can" debate, see Kühler, 2016, 2013b; Collingridge, 1977; Kramer, 2005; Saka, 2000; Vranas, 2007.

Our straightforward intuition tells us that disagreement is the kind of enterprise that presupposes the idea of correctness. As argued in the previous chapter, moral questions or disagreements typically include the idea of correctness. We argue or disagree because we believe there are correct answers; if it were not so, disagreement will not get off the ground in the first place. Nevertheless, Stevenson thinks otherwise. He argues as follow:

> When ethical disagreement is not rooted in disagreement in belief, is there *any* method by which it may be settled? If one means by "method" a *rational* method, then there is no method. But in any case there is a "way" ….where disagreement was due to A's sympathy and B's coldness. Must they end by saying, "Well, it's just a matter of our having different temperaments"? Not necessarily. A, for instance, may try to *change* the temperament of his opponent. He may pour out his enthusiasms in such a moving way - present the sufferings of the poor with such appeal - that he will lead his opponent to see life through different eyes. He may build up, by the contagion of his feelings, an influence which will modify B's temperament, and create in him sympathy for the poor which didn't previously exist. This is often the only way to obtain ethical agreement, if there is any way at all. It is persuasive, not empirical or rational; but that is no reason for neglecting it. (Stevenson, 1937, p. 29)

Of course, Stevenson is right in arguing that we cannot resolve disagreement in attitudes through rational means. The reason is obvious; attitudes are not truth-apt, as we argued above. However, his position is odd, because it is suspicious to assign to any meaningful domain of inquiry such an account of disagreement. Notice that 'agreement' now amounts to emotional outpouring instead of argument; and persuasion instead of rational conviction. First, we might not hesitate in scorning such a method, because of the danger of relapsing into manipulation. The moral domain will look suspicious, if we are to resolve our disagreement through emotional outpouring or persuasion alone. Of course, we are not denying the crucial role of emotions in morality. Second, Stevenson's account of moral disagreement begs the question against the roles of truth-value and rationality in resolving issues in moral discourse. In other words, Stevenson's claim that moral

disagreement is essentially conflict in attitudes seems to fly in the face of how we use ordinarily understand disagreement.

More Difficulties for Emotivism

Generally, expressivism and emotivism in particular claim that moral statements express noncognitive attitudes; and consequently, deny the truth-aptness. As we saw above, this is their semantic position. On the contrary, we believe that moral statements are just like ordinary statements in describing properties of (moral) objects. Thus, moral statements such as 'torture is wrong' can be true or false. The emotivists tell us that the surface features of moral statements differ from ordinary language (given the dynamic functions of moral statements). However, they face the difficulties of explaining the validity and logical features of moral statements. When we utter statements such as, 'torture is wrong' we seem to need a truth-claim, at least, just in case someone denies it. As we argued above, if moral statements are expression of attitudes, we cannot invoke the idea of truth-value; and as such we cannot say that the assertion 'torture is wrong', and its denial 'torture is not wrong', are incompatible.

Consider,

1. It is wrong to torture illegal immigrants.
2. It is wrong for X-government to order its police to torture illegal immigrants.
3. Therefore, it is wrong to torture illegal immigrants and it is wrong for X-government to order its police to torture illegal immigrants.

We can deny either (1) or (2) just in case it is false. However, it is odd, if not totally inconsistent, to accept premises (1) and (2) and reject the conclusion in (3). However, by denying the truth-value of (1) and (2), the emotivists face the difficulty of accounting for the validity of the conclusion reached in (3).[19] If, in other words, we believe that moral statements (such as torture is morally bad) are propositions that admit truth and falsity;

---

[19] The emotivist might, instead of demonstration, decide to offer a causal explanation. That is, he might explain how Anna came to exhibit the emotional response in question. It is still doubtful if such an explanation would supply for the role of language and logical features found in moral discourse. Moreover, it is debatable whether we can say that the emotivist (in the first place) engages in logical demonstrations in the strict sense.

and that their truth-values do not depend on agents; we can explain why torture or suffering is bad, hence, why it is *true* that torturing immigrants is morally wrong. We can as well show *how* we arrived at our *conclusion.* To illustrate,

4. It is a moral fact that torture is wrong and bad

5. Causing immigrants to suffer physical injuries is torture

6. Therefore, torturing immigrants is morally bad.

Furthermore, there is another difficulty, namely the idea of accounting for valid inference. Precisely, it is the challenge of providing plausible semantics for normative sentences in natural languages. It is believed that moral linguistic properties allow us to use predicates such as 'wrong' just as natural predicates (e.g. white). On the contrary, noncognitivism denies the descriptive feature of moral statements. Consequently, moral statements do not have same the semantic force as natural language. However, in order to make sense of moral discourse, any meaningful metaethical theory needs to offer an account of the embeddedness of moral statements in unassertive grammatical contexts. This problem was first raised by Peter Geach (1960, 1965); he, however, attributed to Gottlob Frege's distinction between content and assertoric force. This objection has been since referred to as Frege-Geach Problem. To illustrate how this objection threatens emotivism (of course, it applies to noncognitivism in general, especially the various positions that maintain that moral statements express conative states), let's consider the following:

7. If torture is wrong, then torturing illegal immigrants is wrong.
8. Torture is wrong.
9. Therefore, torturing illegal immigrants is wrong

We can expand it further:

10. If it is wrong to torture illegal immigrants, then it is wrong for X-government to order its police to torture illegal immigrants.
11. It is wrong to torture illegal immigrants.
12. Therefore, it is wrong for X-government to order its police to torture illegal immigrants.

Both premises (7) and (10) come in conditionals; and the moral predicate 'wrong' first appeared in the antecedent in an unasserted form. However, premises (8) and (11) assert something about the wrongness of torture in general, and torturing illegal immigrants in particular, respectively. It is unclear how the emotivist can explain the figuring of moral predicates (like 'wrong') in simple and complex sentences. For him, the meaning of 'wrong' in (8) is different from 'wrong' in (7); and any attempt to construct moral *modus ponens* is most likely to fail in advance. In other words, if moral statements express emotions or ejaculations much like saying 'Ugh!' or 'Boo!', it is difficult to imagine how 'torture is wrong' would look in the if-then statements. In addition, the emotivist will also account for moral predicate (wrong) in the unasserted as well as the asserted contexts: we expect the meaning of the moral predicates across the whole simple and complex contexts. The embedding problem leaves the emotivists in need of some plausible version of expressivism.[20]

Hybrid Expressivism and Mind-Independency

To what extent, if at all, is our moral thought and talk meaningful and objective? Precisely, in giving our attitudes voice as the expressivist claims, can we engage in a meaningful, and an objective moral discourse? With regard to the meaningful aspect, we argued above that expressivism lacks the adequate resources for satisfying both the semantic and logical requirements for moral discourse. But, can the expressivists offer an account of moral objectivity? To answer this question, we will need to distinguish between two expressivist approaches, namely, revisionism and accommodationism. The former is a reaction against British naturalism, the claim that moral statements have descriptive and cognitivist features (e.g. Ayer, Stevenson, Hare); whereas, the latter holds both expressivist and realist-like positions (e.g. Blackburn, Sinclair, Gibbard, Timmons). Given its mind-dependence stance, we can say that the objectivity-talk, on the revisionist analysis, appears not to be possible; after all, to Ayer's mind, moral statements have "no

---

[20] Although Blackburn (1984, 1993) tried to solve the problem, but his position is still highly debated. For more discussions and criticisms of Blackburn's solution see Hale 1986, 1993.

objective validity whatsoever". (Ayer, 1952, p. 108) Accommodationism (or hybrid expressivism as I will refer to it here) claims that expressivism is compatible with moral objectivity – moral mind-independence. In what follows, I however argue that, regardless of its truth-claim, hybrid expressivism does not sufficiently justify our ordinary understanding of moral objectivity.

Hybrid expressivism emerged as a reaction to emotivist semantic and logical deficiencies. As we argued above the emotivist construal of moral judgement makes moral values of things dependent in a way that makes morality susceptible to change according to the agent's attitudes. The goal of hybrid expressivism is, to some extent, to supplement for the role of language in moral practice as well as show that expressivism is, in fact, compatible with moral mind-independency. Our discussion of hybrid expressivism will focus on Simon Blackburn's defence, because his position retains expressivism while arguing that moral objectivity is justifiable on a minimalist framework. His project is basically twofold, namely *projectivism*[21] and *quasi-realism.*

As an expressivist, Blackburn claims that moral judgements are not beliefs that describe moral facts. When we make moral judgements, we express attitudes on situations, actions etc. Adopting Hume's idea, Blackburn argues that moral judgements are projections of attitudes.

> We say that [we] *project* an attitude or habit or other commitment which is not descriptive onto the world, when we speak and think as though there were a property of things which our sayings describe, which we can reason about, know about, be wrong about, and so on. Projecting is what Hume referred to when he talks of "gilding and staining all natural objects with the colours borrowed from internal sentiment", or of the mind "spreading itself on the world." (Blackburn 1984, pp. 170 - 171)

---

[21] The idea of projectivism in philosophy goes back to Hume. In his *Treatise of Human Nature*, he claims that when we engage in talks about external objects, we 'project' or 'spread' our internal sentiments on them. He writes that, "Tis a common observation, that the mind has a great propensity to spread itself on external objects, and to conjoin with them any internal impressions, which they occasion, and which always make their appearance at the same time that these objects discover themselves to the senses" (Hume THN, 167).

On this analysis, we give voice and spread our attitudes *as if* there were things out there. Already, we can anticipate Blackburn's goal, namely, to make attitudes sensitive to situations, people, actions in the world; but, not in the sense of describing things in the world. Do note that it is the agent's attitude projected onto the world; and if judgements were to be so, then projectivism would also face the moral-dependency threats. Blackburn was aware of this threat (this also applies to other hybrid expressivists; see Sinclair 2008; Gibbard 2003; Horgan and Timmons 2009 etc.). If we should spread different attitudes to the same object, we might face the problem of spotting who is in error. Thus, it seems we cannot talk about objectivity in morality if our talk lacks truth-value. This is exactly the hybrid expressivist objective: to integrate truth-value into the expressivist framework. However, to his mind, we do not need to adopt the metaphysical commitment of realism to account for the role of truth in moral talk. "The projective theory intends to ask no more from the world than we know is there [...] It ask no more than this: a natural world, and patterns of reactions to it" (Blackburn 1984, p. 182). Hybrid expressivism is *minimalistic* in its conception of truth.

Blackburn attempts to straddle the divide between expressivism and moral realism, precisely, cognitivism – *Quasi-realism.* Blackburn's position is, without doubt, ambitious because it is an "attempt to have one's cake and eat it too in metaethics" (Street 2011, p. 1). Zangwill notes that "Quasi-realism has enabled the realism debate in moral philosophy to take a new and hitherto unsuspected dialectical turn" (Zangwill, 1994, p. 205). It is attractive, because it seems not only to dodge the talk about whether morality is something, but also it claims to explain the moral semantics. As a deflationary theorist, Blackburn argues that moral judgements, although they express attitudes, are accessible in terms of truth and falsehood. However, they are not robust truth claims; and as such, we should not search for any truth-maker (the sort the realist construes robustly) beyond the statements in which those truth claims occur. To illustrate, according to the minimalist about truth:

I.   *It is true that X if and only if X.*
II.   *'X' is true if and only if X*

The truth or falsehood of (I.) and (II.) lies within the internal features of these statements. Although, it has a 'cognitivist feel', this position rejects any metaphysical claim of substantial properties or facts. This position is consistent with the expressivist claim that moral judgements do not describe moral facts. For hybrid expressivists, if we do away with those facts, we can conceive moral judgements, after all, as truth-apt. On Blackburn's analysis, we can say that,

(II*) It is true that torture is morally wrong if and only if torture is wrong.

However, is minimalism compatible with hybrid expressivism? If the projectivist insists that moral judgements are not beliefs but a projection of attitudes, it is difficult to see how truth figures in an agent's attitudes. Attitudes are conative in nature, and not supported by the psychology of belief, as we argued above. The hybrid expressivist's task is no less challenging, for it wants to justify the truth-aptness of moral judgements on the expressivist framework. Do note that moral judgements express attitudes. Thus, we might still accuse the hybrid expressivist of saying that:

(II**)It is true that X is wrong, because, our projected attitude tells us that it so.

We can translate this claim into,

(II***)It is true that X is wrong, because I think that it is wrong.

If I think X is wrong, then X is wrong

If I disapprove of X, then X is wrong.

The position still makes wrongness of X dependent on the agent's attitude in a way that is objectionably mind-dependent notwithstanding 'it is true' appearing in (II**).

Normally, we do not conceive morality as an output of projected attitudes. But in (II**) and (II***), moral judgement follows the force of the agent's attitudes: If the agent were to project different attitudes onto X, his moral judgement will come out differently. However, we do not expect the wrongness of X to behave in such a manner, lest nothing will be right or wrong in itself. On the minimalist construal, the truth-value of these statements is internal as well as responsive to attitudes. But when we make moral statements, we normally use them to assert something which is outside the statement – something about the object. For example, by asserting that 'torture is morally wrong', the wrongness of torture is supported by an external fact; and truth-value is not dependent or responsive to our belief, thought, attitudes, or opinion. Blackburn was aware of deep mind-dependent consequence of balancing moral values as well truth-apt on attitudes, when he made the following statement: **"**We discover such facts, we do not invent them. It is not because of the way we form sentiments that" [torture is wrong]. "It would be wrong whatever we thought about it" (Blackburn 1984, p. 217). However, the difficulty still remains: is it not counterintuitive to claim that we discover 'objective' facts; and at the same time, deny that they do not figure substantially in moral statements? What use do those facts then have in the truth-value of moral statements?

Blackburn responds to this worry by arguing from the *internal reading:*

The counterfactual 'If we had different attitudes it would not be wrong to kick dogs' expresses the moral view that the feature which makes it wrong to kick dogs is our reaction. But this is an absurd moral view, and not one to which the projectivist has the least inclination. Like anyone else he thinks that what makes it wrong to kick dogs is that it causes them pain. To put it another way: he approves of a moral disposition which, given this belief as an input, yields the reaction of disapproval as an output; he does not approve of one which needs belief about our attitudes as an input in order to yield the same output, and this is all that gets expression in the counterfactual. (Blackburn 1981, p.179)

From the foregoing, an agent's attitudes are divided into first-order sensibility, and second-order judgements. The former corresponds to the initial sensibility of an agent's

56

approval or disapproval toward an action or a situation. In Blackburn's example, it corresponds to our initial reaction of approval or disapproval toward kicking a dog. The second-order judgement does not depend on the content of the former, but on a further element, namely, pain resulting from kicking a dog. Hence, on this analysis of mind-independence, an agent does not approve or disapprove an action or a situation based on his initial attitudes; rather, he links his initial attitude to the content of the second-order attitude (for example, the pain which kicking a dog causes). If this is correct, then Anna did not disapprove the torturing of immigrants based on her initial response to the act but based on the pain that arises from torture. In other words, the proponents of hybrid expressivism want us to understand moral judgement as follows:

> E4: When we judge something as morally good and bad, we tie our initial reactions to elements that are independent of our feelings or attitudes. Precisely, it is our second-order judgement that is expressed in our first-order attitude.

However, does this really capture the notion of mind-independence? Notice that the second-order reaction to the natural element is an attitude, hence an agent's projection onto the object. Since the second-order judgement is an attitude, which is conative in nature, it is not clear how such attitude is supposed to track the facts of the natural elements. Having a second-order attitude toward the first-order attitude of moral judgement does not really save hybrid expressivism from the charge of mind-dependency. This is because; its second-order activities are not structured to track realities. Given that, both orders are essentially attitudes, which are embedded in conative state (and not a belief which admits either truth or falsehood), it is difficult to see how these attitudes which admit gradations are supposed to capture the natural elements of things and at the same time relate them in terms of truth and falsehood.

Furthermore, agents can possess different reactions to natural elements even on the second-order attitudes. It is possible that Anna and Ted would have different attitudinal reactions toward pain, which might result in divergent moral judgements. If this case is possible, then it seems that moral claims are still dependent on the attitudinal stances of the agents. In addition, such a possibility has an effect on the formation of moral sensibilities. In reacting to the belief-like statement: 'it is wrong to torture

immigrants'; Anna forms attitudes towards the action. But then again we seem to require some sort of standard by which we check if the attitude projected captures the wrongness of torture. Given that we can possess different attitudes, we need, at least, a framework by which we can separate appropriate attitude from inappropriate attitude. Assuming the hybrid expressivist claims that our attitudes are equal, then subjectivism would threaten again. We need to endorse some attitudes and renounce others. Hence, to make the objectivity-claim meaningful, such a framework should apply not only to Ann, but also to everyone. Blackburn's argument from idealization (that is, appropriate attitudes should be formed on the basis of "sensitivity, maturity, imagination and coherence" (Blackburn 1998, p. 318) is insufficient because the framework is relative to individuals. Do note that sort some of relativist mind-dependence threatens here as well. Warenski argues thus,

> I think the expressivist can give an account of the *conceptual status* of moral mind-independence on an internal reading, but when this internal reading is coupled with an internal reading of *correctness or appropriateness of moral judgment*, it cannot fully secure mind-independence. When correctness or appropriateness of moral judgment is given an internal reading, the conceptual status claim can assume that appropriate moral values will reflect moral framework principles that are stable on reflection and free of readily-correctable faults of reason or evidence. This will be sufficient to secure a partial mind-independence: moral judgments would not be immediately responsive to occurrent attitudes in the way that the mind-dependence conditionals suggest. However, moral judgments would still be responsive to the attitudes that underlie moral framework principles, and these attitudes may or may not generate principles that coincide with those moral commitments that the quasi-realist wants to preserve. (Warenski, 2014, p.866)

There are, at least, two main options open to the hybrid expressivist: either he goes for the convergence of idealized attitudes – common framework of idealization; or he adopts an external framework independent of the individuals' attitudes. Both approaches come with heavy cost: to the first option, Blackburn has to offer an account of how the individual idealized attitude is a "set of desire that was maximally informed and coherent

and unified" (Smith, 2003, p. 312). To the second, he has to adopt not just a realist-like position, but a form of realism where the existence of moral facts is possible. Nevertheless, it might still be argued that the second-order reaction is independent of an agent's first-order attitudes. In other words, the hybrid expressivist can argue that the agent's attitude on the second-order platform objectively expresses why it is morally wrong to kick a dog, cat or goat. Nevertheless, this line of argumentation will put hybrid expressivism in a tension. This is because, on the one hand, it rejects the claim that moral judgements are beliefs; and on the other hand, it accepts that it is the natural elements of things or actions that determine values of moral claims. In this case, he saddles the burden of explaining the nature of such natural elements. That is, he has to answer whether such natural elements (for example pain) are part of the fabric of the world. If yes, in what ways, if at all, do they differ ontologically from other natural facts of the world? Notice that if they do not differ from other natural facts, then the hybrid expressivist construal of mind-independency will make sense if only he adopts some ontological aspects of moral realism, precisely the naturalist version of moral realism. Nevertheless, accepting this option forces the hybrid expressivist to accept as well the robust truth claim: thereby espousing the sort of metaphysical commitments that he abhors.[22]

2.3. Conclusion

According to noncognitivism, moral questions are not about anything. When we engage in moral discourse there is nothing in particular which our thoughts and talks refer to, or reflect, because such properties or facts do not exist. Precisely, we encountered the moral expressivist claims, namely that moral judgements do not describe things in the world, rather they express conative attitudes; and that the psychological state supporting moral judgements is noncognitive. I argued that expressivism makes morality objectionably incoherent and heavily dependent on an agent's attitudes. I focused on two manifestations of expressivism: emotivism and hybrid expressivism. I argued that emotivist claim seems

---

[22] Nevertheless, moral mind-independence is possible on the hybrid expressivist framework but via the external reading. For more discussion see Zangwill, 1994; Jenskins, 2005; Warenski, 2014.

to be incoherent: it cannot deny the existence of moral facts and still makes claim about and in the domain which it rejects. But more fundamentally, I argued that by balancing morality on an agent's attitudes or emotions, the emotivist faces the difficulties of explaining and communicating morality objectively; and he also fails to explain moral disagreement, as well as the semantic and logical features of moral discourse. Furthermore, I argued that hybrid expressivist construal of moral objectivity is insufficient. It does not fully escape the moral dependency objection, given the role of attitudes in moral judgements. Precisely, Blackburn's attempt to justify moral mind-independence via internal reading technique requires an external reading; however, such resources are available to moral realism, but not to expressivism. Thus, in the next chapter, I will argue that in providing an account of the theoretical feature, we are better off adopting a position that explains how moral values of things are not dependent or caused by moral agents' attitudes. Moral realism seems to be a better alternative; hence, we have reasons to abandon moral expressivism in its two versions. Moral realism does "justice to moral thought" (Zangwill 1994, p. 205).

Chapter Three

Moral Realism

> We use our imagination not to escape the world but to join it…The value concepts are …patently tied on to the world, they are stretched as it were between the truth-seeking mind and the world, they are not moving about on their own as adjusts of the personal will. The authority of morals is the authority of truth, that is, of reality.
>
> ***Iris Murdoch, 'The Sovereignty of Good', 1970***

3.0. Introduction

The objective of this chapter is twofold: First, to characterize moral realism, and second, to argue that moral realism best explains the idea of objectivity of morality. I start by outlining the defining features of moral realism: These features equip us with the tools for separating moral realism from other forms of realism as well as other versions of cognitivism. I argue that moral realism entails moral cognitivism but not vice versa. Hence, in order to effectively characterise moral realism, our criteria must go beyond the features of cognitivism. I, then, describe the nature, namely, the metaphysical assumptions underlying moral realism. I then go to defend a nonreductive naturalist version of moral realism relevant for the purposes of this project.

3.1.   What is Moral Realism?

Realism is characterized by its metaphysical, epistemological and semantic claims;[23] and it comes in different flavours. Hence, the realist-claim in some domain or subject-matter, 'S', might differ from the claims in other domains or subject-matters, 'M', and 'Y'. For example, it is possible for someone to be a mathematical realist and an aesthetic irrealist.

---

[23] It is good to note that realism is used both in the ordinary sense and technical sense. In this work, I employ it in its technical sense (technically, realism appears in mathematics, sciences, psychology, law, art etc.). However, our focus is on philosophical realism (in its philosophical manifestation, realism appears in the domains of philosophy of mind, aesthetics etc.), precisely on moral realism.

In addition, it is possible to be a moral realist and a mathematical irrealist. Given their distinctive characters, it is in order to map off these different domains or subject-matters of realism. Geoffrey Sayre-McCord offers one of such map-off models. In his *Essays on Moral Realism*, he writes:

> The map I offer is, I believe, one that works not just for the moral realism debate, but also for all other debates concerning realism. Moved from realist to realist debate, the same map will serve to demarcate in each the issues that separate realists from anti-realists. Unlike most maps, which elaborately detail one area but are useless in the next, this map serves as a guide to realism wherever it occurs. (Sayre-McCord 1988, p. 5)

One of the advantages of the model is that it enables us not only to map off moral realism from other forms of realism (that is, it helps us to characterize the defining features of moral realism), but also it allows us to separate moral realism from anti-realism. Deborah C. Smith observes that:

> With such a map in hand, identifying the moral realist should be relatively easy; any moral theory that falls within the territory marked off by the map is a realist moral theory. If the criteria are not successful, we may gain some insight into the nature of moral realism through examining the reasons for their failure. (Smith 2001, p. 24)

Further, Sayre-McCord offers two criteria for categorizing various forms of realist claims:

1. The claims in question, when literally construed, are literally true or false.

2. Some of the claims in question are literally true.

The first criterion requires that claims made in some domains, 'S', 'Y' and 'M', be truth-apt; while the second criterion holds that, at least, some of the claims of these domains are in fact true.[24] Since our interest here is on the moral domain, it is largely argued by

---

[24] The second criterion separates (moral) realism from Error Theory, generally a form of (moral) cognitivism which holds that all (moral) propositions are false. In other words, they cannot be true. We shall return this position later.

the moral realists that moral claims are truth-apt and not all of them are false. At least, some of them are true (Boyd 1988; McNaughton 1988; Brink 1986, 1989; Smith 1994; Schaber 1997). However, Sayre-McCord's model is silent about how the truth-value of said claims is to be construed. That is, it does not tell us whether the moral facts underlying the truth claims are to be understood robustly or minimally. In fact, it is neutral about the existence of such facts. McCord left it so for a reason:

> Independence from the mental may be a plausible requirement for realism when we're talking about macrophysical objects but not when it comes to realism in psychology (psychological facts won't be independent of the mental); bivalence might go hand in hand with realism in mathematics, but realism in other areas seems perfectly compatible with acknowledging that some of our predicates are vague and have indeterminate extensions; and existence may be crucial to realism about scientific entities (since claims concerning such entities are true only if the entities exist) but not realism about scientific laws (that makes no existence claims). (Sayre-McCord 1988, p. 6)

It is true that such attribution would amount to precluding some realist positions. However, inasmuch as philosophical realism (moral realism to be more precise) is concerned, satisfying the above criteria does not make a theory a realist position. I will show that construing moral realism on these criteria place us under the threat of qualifying non-realist (and even anti-realist) theories as moral realism.

3.2. Not Morally Realistic Enough

On Sayre-McCord's construal, 'S-Type' theory is realist because its claims are truth-apt and some of the 'S'-statements are literally true. The strength of this construal is that it provides us the tool for characterizing cognitivist positions. However, it does not sufficiently help us in defining realism. Hence, the claim that his map is effective in all domains of realism is not quite true: At least, characterizing a position solely on the

grounds of cognitivism does not necessarily qualify such position as moral realism. Moral realism entails moral cognitivism but not vice versa.[25] To illustrate, let us consider a moral theory which holds that 'Peter ought not to shoplift'. The statement expresses a state of affairs, such that it is true that shoplifting is morally bad. By this token, it would be false to think that Peter, after all, can shoplift in shop B, but not in shop A. Notice that we can argue that our moral claim is a truth-apt. Thus, it fulfils criterion (1). Further, suppose we have grounds to believe that the statement: 'shoplifting is morally bad' is, in fact, true. With this, we have fulfilled criterion (2); and going by Sayre-McCord's model, this moral theory can be said to be a form of realism. The truth-claim of a given domain does not necessarily make it a realist position (thus, more features are required). If this were to be so, Terry Horgan and Mark Timmons' metaethical theory would be classified as a form of moral realism, because it satisfies both criteria. But is it moral realism?

In their work *Morality without Moral Facts*, they set out to propose a position that seems to capture our ordinary moral thought and discourse. They argue that moral beliefs possess phenomenology of an occurrent belief, namely *what-it-is-likeness typically*, which involves:

(1) Psychologically 'coming down' on some issue, in a way that

(2) Classifies (sometimes spontaneously) some 'object' of focus as falling under some category, where one's classificatory coming down is experienced

(3) As involuntary,

(4) As a cognitive response to some sort of consideration that is experienced (perhaps peripherally in consciousness) as being a sufficient reason for categorizing as one does, and

(5) As a judgment that is apt for assertion and hence is naturally expressible in public language by a sentence in the declarative mood. (Horgan and Timmons 2009, p. 263)

---

[25] This is true for other forms of realism, especially the philosophical realism where cognitivism is not insufficient for characterizing a theory as realism.

Although moral beliefs share these key generic phenomenological features with other non-moral beliefs, Horgan and Timmons argue that they do not have the overall descriptive content. This is because mental states underlying moral beliefs and non-moral beliefs are distinct psychological *sui generis* contents. Hence, mental states of moral beliefs are irreducible to those of non-moral beliefs. In addition, the mental states supporting moral beliefs do not have the fittingness of representing the world in a certain way or as it might be, because "To construe moral beliefs in this manner is to mistakenly assimilate them to descriptive beliefs, i.e. to is-commitments. Rather, an ought-commitment is a distinct kind of mental affirmation vis-à-vis a core descriptive content"(Horgan and Timmons 2009, p. 271).

However, they strongly contend that non-descriptive contents of moral beliefs are compatible with truth and assertoric features of non-moral beliefs. One of the implications of their claims is that we can assign truth-values to moral beliefs without having to describe the world as it is or might be. In other words, moral beliefs can be genuine beliefs, which can be true or false assertions. Nevertheless, we do not need such things as in-the-world moral facts, namely, truth-makers (Horgan and Timmons 2009, p. 275). For them, when we say that moral judgements and statements are true or false, we are engaging in moral semantic appraisals. That is, "appraisals in which semantic evaluation are 'fused' with moral evaluation. These truth ascriptions thus are not descriptive, because the overall content of the first-order judgments and utterances to which they are applied is not descriptive" (Horgan and Timmons 2009, p. 275).

Like Blackburn's construal of truth-value we saw in the previous chapter, Horgan and Timmons argue that truth ascriptions are supposed to function within the framework of minimalist Schema T: *'S' is true if and only if S*, where we can substitute 'S' for any (moral) declarative sentence. In their cognitivist expressivist model, 'X is wrong' is true 'if and only if X is wrong'. Thus, the moral judgement: 'Peter ought not to shoplift' translates into: Peter ought not to shoplift is true if and only if Peter ought not to shoplift. The statement is belief-like. On this analysis, the 'is-commitment' belief and 'ought-commitment' belief express some possible state of affairs. However, unlike the former, the latter does not refer to any fact in the world, because such 'facts of ought' or precisely

facts of morality do not exist anywhere in the world. In their example, both 'John gave back to Mary the money he owed her' and 'John ought to give back to Mary the money he owes her' fulfil the requirements of genuine beliefs.[26]

> Both kinds of commitments state are beliefs since they exhibit certain generic features that characteristic of beliefs…both sorts of commitment state have the grammatical and logical trappings of genuine beliefs: in thought and language the contents of such states are declarative, and they can figure as constituents in logically complex judgements as in 'Either John has paid what he owes to Mary or he ought to do so'. As such, ought-commitments can figure in logical inferences. (Horgan and Timmons 2009, p. 232)

Notice that if statements such as 'Peter ought not to shoplift' or 'John ought to give back to Mary the money he owes her' are genuine beliefs, then they are accessible in terms of truth and falsity; and as genuine beliefs, some of the claims they express are true. Based on Sayre-McCord's model, we can qualify Horgan and Mark Timmons' position as moral realism. However, fulfilling those two criteria leads one to moral cognitivism, but not necessarily to moral realism. In other words, while we can regard this position as a form of moral cognitivism, we cannot classify it as moral realism. This is because they reject the kind of facts that make realist moral judgements or beliefs true. I regard Sayre-McCord's criteria as a model for marking off cognitivism, but not necessarily for realism. Hence, we can reformulate (2) thus,

> 2b. Some of the claims in question are literally true, because there are facts (x), (y), (z) which they refer to.

Given this, we can mark off metaethical positions that espouse the above criteria, but deny the existence of moral facts such as (x), (y), (z) as irrealism. The insufficiency of criteria (1) and (2) applies as well to the sort of metaethical theory defended by John

---

[26] It goes without saying that Sayre-McCord's model seems to make moral subjectivism a moral realist position. Consider a subjectivist saying, X is wrong where this means I disapprove of X. Given that 'X is wrong' is seen as a belief, it appears that subjectivism is moral realism. But the fact supporting the statement is different from the moral realist fact.

Mackie, namely Error Theory. In *Inventing Right and Wrong* (1977), he holds that moral judgements are both objective and normative in character. That is, they prescribe objective action-guiding values and principles that are parts of our natural world. If we stop here, we might classify Mackie's position as a realist position. Notice that, unlike Horgan and Timmons' position, Mackie's seems to hold that statements such as 'Peter ought not to shoplift' does not just affirm a 'metalinguistic first-order moral claim in question' (Horgan and Timmons 2009, p. 234), but they seem to be referring to some sort of objective value in the world. On this analysis, Mackie appears to be espousing naturalist ontological commitment; hence, we might be tempted to call his position realism. However, Mackie strongly argues that such purported objective facts and properties we employ in ordinary moral thought and discourse do not in fact exist. To his mind, there are no objective facts, the type that purportedly describe and make moral judgements true. If there were to be such facts, given their normative character, their nature would be ontologically queer. Thus, since such facts do not exist, all our moral judgements are false,[27] because they were built on error. At this juncture, we have to strip Mackie's view of the realist character we previously assigned to it, because it rejects the type of facts that make realist moral judgements true.

Nevertheless, there are some lessons we can learn from both positions discussed above. They tell us not only how we are to demarcate cognitivism from realism (that is, views that are cognitivist, but devoid of moral facts), but also they offer us some guides on how we are to characterize the constitutive features of moral realism. Based on this, we can improve on Sayre-McCord's criteria. Notice that as a consequence of not specifying the ontological assumptions of his model, Sayre-McCord seems to blur the distinction between the first-order moral claim and second-order moral claim. The former position holds moral propositions as truth-apt. However, it denies that the truth-maker of moral propositions is based on moral facts and the like. The latter position claims that moral propositions can be true or false because of the existence of moral facts. This distinction is very crucial, because an irrealist might, as well, hold that the moral

---

[27] Given that moral realism holds that some of the moral claims are literally true, Mackie's assumption, namely, all moral judgements are false, safely puts him on the irrealist camp on the Sayre-McCord's Model.

propositions are literally true or false, where the truth-value in question is understood on the first-order reading: This is exactly the position Horgan and Timmons defended. Unlike Mackie's Error Theory, they do not think that all moral propositions are based on errors. However, although they conceive moral propositions as genuine beliefs, (which are, in turn, truth-apt), the truth-aptness of moral claims are not determined by moral facts. To their mind, "when one predicates the truth of a moral statement, one is engaged in an act of affirming "metalinguistically" the first-order moral claim in question (that is, affirming first-order moral judgment expressed by the statement one is calling true). Such an affirmation, done metalinguistically by employing the semantic concept of truth, is a morally engaged "fusion" of semantic and moral appraisal" (Horgan and Timmons 2009, p.234). Alternatively, separating the first-order moral claim and second-order claim of (2) is important because it enables us to ascertain if the type of facts that make moral judgement true is compliant with realist conception of moral facts. In order to mark off moral realism from other forms of realism and irrealism, it is insufficient to posit (1), (2) and (2b) criteria. Also, we have to include in our model the first-order and second-order conceptions of moral claims; and verify if the ontological requirement, that is, if the moral facts are the type that moral realism upholds. In sum, it is the belief that 'moral claims are not only truth-apt but also, that the truth-maker of such moral claims is moral facts' that essentially drives moral realism. We shall see shortly that a truth-claim supported by moral facts enables us to provide a plausible account of moral mind-independence. Precisely, ontological claim of moral realism is robust enough to avoid the threat of making morality objectionably dependent on agents' attitudes, emotions, feelings, interests, preferences, etc.

3.3. Moral Objectivity qua Mind-Independency

Moral realism conceives facts as objective facts of morality. The idea of objectivity is tied to some sort of metaphysical claims about the existence and nature of entities, occurrences or relations. The objectivity-claim in some domains, t, touches on the question of an external world, a sort of world in which the agents who talk about objectivity are part

of.[28] Objectivity, if only loosely, is a way we refer to objects or realities that lie outside of us. Consider, for example, Beverly McLoughland saying,

> I usually write at the kitchen table, where I sit facing the living room window. When I look up from my writing *I can see the woods.*[29]

Suppose Beverly can also *see* squirrels jumping off the trees; and beside her, she can *see* her dog lying on its back. This activity involves objects outside of Beverly, hence, objects in the world. By saying, for example, the squirrels are jumping off the trees; she is saying *something* about certain objects – squirrels and trees – which are in the world. Notice that there is a sort of relationship between Beverly and these objects. From a realist perspective, it is believed that such an activity presupposes the existence of the said objects independent of any conscious awareness, perception, or discovery. In other words, on this view, the squirrels as well as the trees, will continue to exist independent of Beverly (or anybody else) *seeing or thinking of* them. Thus, when we say that something is objective, we talk about our relationship with the external world, namely, we talk about realities insofar as they are independent of our personal preferences, thoughts, opinions, and projections of attitudes.

First, notice that the idea of independency is crucial to our objectivity claim. Very roughly, independency is the stance of standing back from personal attitudes, opinions, or beliefs. This fact immediately points to a relationship between the agents and things in the world. At this juncture, reality is used in a loose sense to cover those sorts of things

---

[28]The objectivity-talk spreads across various subject-matters of ordinary practices. Our thought and language operate in a way that presupposes the existence, or at least, possible existence of objective entities out there. However, I think that our objectivity-talk would only make sense if there are beings that are capable of knowing and talking about such an external world. This concern is also related to the issue of truth-value of facts: "If we imagine a world of mere matter, there would be no room for falsehood in such a world, and although it would contain what may be called 'facts', it would not contain any truths, in the sense in which truths are things of the same kind as falsehood" (Russell 1912, p. 170). Of course, this claim is not to be understood to mean that the *existence* of entities or truth-values of facts in question is not possible in the absence of such minds. However, human beings, at least, appear to be the only beings capable of strictly engaging in talks about the objective status of realities as it is in the world, a world which they are part of as well. In a strict sense, stones, trees, mountain, water or lions do not engage in such talks in a way that is relevant to domains of human inquiry.

[29]The emphasis is mine. Beverly McLoughland is an author and a poet.

that *really* exist. By coming in contact with realities, the agent enters into a sort of *relational encounter* with them. This encounter presupposes an existence as well as a link between the 'I' (subject) and the 'Other' (object), whereby the existence of the latter does not depend on the former and vice versa. For example, Beverly stands in a relational encounter with the objects - squirrels on the trees, dog, and woods. The existence of these objects is independent of her mind, and they are not causally dependent on her thoughts or projections of feelings. Hence, one of the main elements of objectivity is the independent existence of entities. If only roughly, this is the ontological assumption underlying philosophical realism in general and moral realism in particular.

Moral realism is essentially construed on the belief that objective facts and properties of morality exist; and their existence and nature do not depend on our opinion, feeling, thoughts or evidence of them. In other words, moral realists believe that there are objective moral facts in the world; and as Murdoch notes "they are stretched as it were between the truth-seeking mind and the world" (Murdoch 1970, p. 90). The task of moral agents is not that of creating them, but discovering these independent facts of morality. Thus, to the realist mind, the kinds of facts that make moral judgements true are *objective*, precisely *mind-independent facts* of morality.

To recap, there are, thus far, three defining features of moral realism, namely:

1.   Moral claims (such as judgements and statements) are capable of being true or false.

2b.   Some of the claims in question are literally true, because there are facts (x), (y), (z) which they refer to.

3.   There are objective moral facts whose existence is mind-independent.

But, how is the moral realist mind-independency to be understood? I shall begin by, briefly, tracing the notion of moral mind-independency back to Plato; and at least, a passage quickly comes to mind, namely the *Euthyphro*. In the dialogue, Plato's Socrates asks his interlocutor, Euthyphro:

Is the pious loved by the gods because it is pious?

Or is it pious because it is loved by the gods? (10a)

Apart from the classical dilemma he posits,[30] what Socrates asks, in other words: "Is something good because we favor it? Or, do we favor it because it's good?" (Railton 2006, p. 201; Miller 2013, p.7) He raises the mind-independency versus mind-dependency issue. He asks: Are the facts that make our moral judgements good or bad dependent on our belief, perceptions, thoughts or evidence about them? Or are they independent of our belief, perceptions, thoughts or evidence about them? Moral realism upholds the latter position, whereas irrealism defends the former. In what follows, we shall consider how the moral realists construe moral mind-independency: This exercise is crucial because it further specifies a defining feature of moral realism. In other words, the three defining features listed above do not tell us about the nature of the moral realist's construal of moral independence.

Weak Moral Mind-Independency

Moral facts can be construed weakly, that is, they are independent of given individual agent's beliefs, perceptions, thoughts or evidence about them. Kramer writes:

> Sometimes when theorists affirm that mind-independence of certain matter, they are simply indicating that the facts of those matters transcend the beliefs or attitudes of any given individual. They mean to allow that those facts are derivative of the beliefs or attitudes that are shared by individuals who interact as a group. These theorists contend that, although no individual's view are decisive in ordaining what is actually the case about the matters in question, the understandings which individuals share in their interactions as a group are indeed so decisive. (Kramer 2009, p.24)

---

[30] On the one hand, if we claim that the pious is pious because the gods loved it, then it would seem that anything the gods say is pious would be so. On the other hand, if we say that the pious is loved by the gods because it is pious, then piousness would not only be independent of, but also above and beyond the gods. Hence the gods would only be describing piousness of the pious. In other words, the gods would not be almighty. This dilemma replicates in the *god and goodness argument.*

The emphasis is on any given agent. The theorist of this form of mind-independence would argue that the facts that make our moral judgements true or false, good or bad, are independent of any agent's belief, perceptions, thoughts or evidence about them. However, groups of individuals can construct or create such facts. If mind-independence is conceived as not dependent on an individual, then this position, by fulfilling the three criteria qualifies to be called moral realism. Nonetheless, this construal fails the moral realist test, because it conceives moral facts as artefacts (for solving moral problems[31]), which are created or constructed by (interactions of) individuals. For example, the chair and the kitchen table, which Beverly usually uses, are constructed, but their continued existence does not depend on the mind of its maker. Similarly, the proponents of weak moral mind-independence claim that moral facts are constructed, hence, causally dependent on rational agents. Moral independence is weakly created when the constructed moral facts are not products of any given individual mind. Once constructed, the continued existence of these facts does not depend on any given mind or that of group of individuals.[32] Kramer rightly notes that such an interaction must not necessarily involve the convergence or agreement of all the members of a given group, however, the "convergence among a preponderance of a group's members will be sufficient to group the existence or establish the nature of some weakly mind-independent phenomenon" (Kramer 2009, p.25).

Suppose, for example, that we rationally converge that 'S' is a moral fact. On the weak mind-independent construal, the continued existence of 'S' transcends minds or thoughts of agents. Suppose that suddenly all the agents that constructed 'S' disappear,

---

[31] So according to constructivism, normative concepts are not the names of objects or of facts or of the components of facts that we encounter in the world. They are the names of solutions of problems, problems to which we give names to mark them out as objects for practical thought.

[32] Korsgaard observes that what separates moral realism from moral constructivism is the role of procedure (hence, procedural realism) in the construction of moral facts. She writes: "The procedural moral realist thinks that there are answers to moral questions *because* there are correct procedures for arriving at them. But the substantive moral realist thinks that there are correct procedures for answering moral questions *because* there are moral truths or facts that exist independently of those procedures and which those procedures track. Substantive realism conceives the procedures for answering normative questions as ways of *finding out* about a certain part of the world, the normative part" (Korsgaard 1992: 38 see also p.36).

'S' would continue to exist.[33] This implies that neither does the continued existence of 'S' depend on the presence of its creators nor on their observation of 'S'. Notice that moral facts would continue to exist even if none of their creators observes them. Hence, moral facts, even on this construal, are observationally mind-independent. "Something is observationally mind-independent if and only if its nature (comprising its form and its substance and its sheer existence) does not hinge on what any observer takes the nature to be" (Kramer 2009, p.25). Regardless of the surface fulfilment of the features listed above, we cannot strictly call any metaethical theory that espouse weak mind-independence a moral realist position, because its moral facts, unlike those of the moral realists, are causally constructed, hence, to some extent dependent on agents.

Apart from failing the moral realist test, such positions are problematic on a number of grounds. First, imagine the cost of bringing all the members of a group to converge on what would count for and against moral facts. Given the nature and character of various groups of agents, it appears difficult to bring all members to a convergence point. In some groups, it might be difficult to even bring all the members to participate in the first place. On this ground, one falls back to the majority of the group in order to achieve such a convergence. Notice that the force of such moral facts (and their corresponding instantiations in moral principles) would heavily rely on some sort of expectation. That is, an expectation that *all* the members of the group (both the majority and minority groups respectively) would adhere to them. The ground for such an expectation might be rationality. It might be argued that since the moral principle (that the majority converged at) is rational, they apply to all rational agents as such. Thus, it is expected that all agents adhere to it. However, given that there is no conclusive model of rationality, this argument fails (we shall return to this issue in chapter 4). It is possible that the minority has legitimate reason not to adhere to the majority's reason. It might as well be the case that the majority camp is in error.

---

[33] Such entities might continue to exist, but if their existence is closely and significantly tied to their utility, then their continued existence would be lame because there will be no agent to utilize their usefulness. In other words, the practical usefulness of the constructivist moral facts would only make sense if there are minds that construe them as practical solutions to moral problems.

Second, in order to avoid biased or prejudiced outcomes, the moral constructivists argue that rational agents or participants of rational deliberations should become *ideal observers*. In other words, these ideal observers enter the process of moral deliberation by taking what Rawls calls the original position:

> The parties in the original position do not agree on what the moral facts are, as if there already were such facts. It is not that, being situated impartially, they have a clear and undistorted view of a prior and independent moral order. Rather (for constructivism), there is no such order; and therefore no such facts *apart from the procedure of construction as a whole*; the facts are identified by the principles that result [from the procedure]. (Rawls 1999, p. 354)

I doubt very much how such a posture is possible given that such a process of moral deliberations relies on some idea of procedures. It raises disturbing questions of justifying a chosen position or procedure; providing reasons why a given procedure is preferred over others; specifying why the original position is to be set up in a given way in the first place; and securing the legitimacy of the final justifications of the outcomes. Recall our 'S' moral fact; suppose it is a result of an ideal procedure of moral deliberation. There are features which 'S' is supposed to exhibit to be an impartial and undistorted outcome of ideal process of moral deliberation. However, if there are not some kind of pre-fixed or prior, at least logical or rational, standards on which to test 'S', it would be difficult to determine if it fulfils the requirements of being a genuine basic moral fact or principle. This is problematic on two grounds. On the one hand, if the constructivist denies the existence of such a fixed and independent system, then he begs a question against the rational requirements of determinacy: There seems to be some sort of independent determinacy requirements (out there regardless of the model in question) which rational deliberations (including the moral ones) strive to capture. If the constructivist, on the other hand, accepts such a requirement, then he has to explain away how the ideal judges are supposed to mute such pre-fixed requirements while entering the ideal procedure of construction. Further, suppose, for the sake of argument, that there are no such prior or fixed requirements. By merely assuming that they do not exist, the constructivist is not only begging the question against what the realist accepts, but as well, the clause 'apart

from' in Rawl's position above transforms the procedure of construction into some sort of independent and fixed requirement necessary for ideal moral deliberation. In other words, the constructivist is telling us that there are no such facts (of the moral realist), but there is another objective and fixed fact (that is independent of rational agents) namely, 'the procedure of construction'.[34] On similar count, Horgan and Timmons argue that the constructivist is caught up in a dilemma:

> That is, in characterizing the conditions of ideal deliberation, if the constructivist appeals to the relatively uncontroversial formal and substantive platitudes associated with the concept of being an ideal moral judge, the result will be that there will not be enough constraints on what counts as 'ideal' deliberation to yield determinate moral norms. So, to narrow the field of competitors, the constructivist is going to have to build in some substantive moral assumptions. What will guide the constructivist here? It looks as if the constructivist will have to allow ideal deliberators to fall back on their own deepest moral convictions. (Horgan and Timmons 2009, p. 228)

By taking the first horn of the dilemma, moral deliberation would be unable to produce any outcome or moral norms. If the constructivist takes the second horn, then he would have to accept the fact that ideal agents will converge differently on what counts as 'S'. "But…if a particular ideal deliberator happens to start the deliberative process with deep moral convictions" (Horgan and Timmons 2009, p. 228), then different outcomes that fulfil the requirements of 'S' would certainly emerge. In sum, apart from the fact that the nature of moral facts construed on weak mind-independent platform is not compatible

---

[34] Rawls, indeed, acknowledges the problem and tries to address it in terms of an underlying wide reflective equilibrium, the notion of a hypothetical state of affairs arrived at in the course of trying to justify the original position by resolving expected inconsistencies between our considered judgments and the principles yielded by a candidate description of the initial situation. Although this technique is intended to serve as a justification for the design of the original position and, thus, the procedure of construction, it is not without criticisms. For example, Kelly and McGrath (2010), and Arras (2007) contend that Rawls' reflective equilibrium does not secure the convergence claim. For other criticisms of the technique see for example, Cohen 2007, 2008; Brandt 1979, 1990; Hare 1973; McMahan 2000).

with the nature and existence of moral realist facts, the problems we identified above give us reasons not to construe moral facts in that manner.

Robust Moral Mind-Independency

On the moral realist construal, moral facts are conceived as robust mind-independent entities. Just as natural objects (such as stones, trees, squirrels etc.), so also the moral realists largely think that moral facts are part of the world and independent from any mind. The mind-independency in question here is about the existence and nature of moral facts. Kramer writes that "something is existentially mind-independent if and only if its occurrence or continued existence does not presuppose the existence of any mind(s) and the occurrence of mental activity" (Kramer ibid. p.25). On this analysis, there are two forms of existence in question, namely what I refer to as 'original existence' and 'continued existence'. The former implies that moral facts do not owe their existence, that is, their *coming-to-beingness* to any individual mind or group of minds; whereas the latter maintains that the continued existence, that is, the *continued beingness* of moral facts is not dependent on any individual mind or group of minds. It is in the former form of existence that the robust mind-independent construal differs from the weak mind-independent position. Both positions believe that the continued beingness of moral facts is not dependent on any mind(s).

However, one of the challenges facing the defenders of robust moral mind-independence is that of accounting for the coming-to-beingness and nature of moral facts. In other words, moral realism carries the cost of answering the following questions: Are moral facts and properties radically different from other non-moral facts, hence *sui generis* facts? Where do they acquire their normativity from? Or are moral facts and properties natural facts of the world? It goes without saying that these questions are crucial to the credibility of moral realism. To this end, Peter Schaber writes:

> Diese Fragen muss der moralische Realist beantworten. Denn solange kein plausibler Begriff, moralischer Tatsachen vorliegt, ist die These, moralische Urteile seien wahrheitfähig, eine bloße Behauptung… Solange dies nicht geleistet

ist, bleibt der moralische Realismus eine obskure Position.[35] (Schaber 1997, p. 17)

We shall return to the issues raised by Schaber in the next section, but before then let us further consider other features of robust mind-independence.

Recall the question implied in Socrates' concern, namely "Is something good because we favor it? Or, do we favor it because it's good?" Given the conception of robust mind-independence, the moral realist answers that we favour something because it is good (but this is not sufficient because the defenders of weak mind-independence hold the same position). In other words, moral realism is not to be strictly understood in terms of (3): 'There are objective moral facts whose their existence is mind-independent'. When defined on this view, there is a looming threat of mischaracterization: This is because the concept of existence can mean original existence as well as continued existence. Hence, the kind of facts that make moral judgements, on the moral realist view, are objective and mind-independent moral facts (where moral mind-independency is understood robustly). That is, the existence and the continued existence of such facts in question do not depend on our beliefs, perceptions, thoughts or evidence about them. Thus, we can modify (3) as follows:

(3b) There are objective moral facts; and their existence and continued existence are mind-independent.

Nevertheless, there are some questions still unanswered: Construing moral realism just from the perspective of mind-independence does not tell us about the nature of these objective moral facts. As we noted earlier, it is not clear whether these moral facts are non-natural or natural facts. Apart from the concern raised by Schaber above, investigating into the nature of these moral facts is crucial because it enables us to clarify the metaphysical commitments of moral realism. As we shall see in a moment, the metaphysical justificatory commitment that a naturalist shoulders differs from that of a non-naturalist. And it is without a doubt that the success of moral realism to a large extent

---

[35] "Moral realism must answer these questions. For as long as there is no plausible concept which justifies moral facts, the claim that moral judgements are truth-apt would be a mere assertion … As long as this justification is lacking, moral realism will be an obscure position" (Schaber 1997, p. 17 translation is mine).

depends on a plausible construal of the nature of moral facts and properties. I think that one of the main drivers of moral scepticism, precisely the attack on moral realism, lies on the construal of the nature of moral facts.

3.4. One Castle, Two Towers: Naturalism and Non-Naturalism

Up to this point, we have characterized moral realism as a metaphysical neutral position, such that any moral theory that fulfils the defining features we considered above might qualify as a moral realist position.[36] There are at least two main ways of understanding the nature of moral facts within the moral realist tradition, namely *Moral Non-Naturalism* and *Moral Naturalism*. Moral non-naturalism conceives moral properties, facts, and relations as neither identical nor reducible to natural entities. On this view, moral properties and facts are *sui generis* entities and not the subject-matter of empirical investigations. According to non-naturalists, moral predicates that figure in moral judgements refer robustly to non-natural moral properties. In other words, the moral fact that makes the moral judgement: 'it is morally wrong to torture people for pleasure' exists mind-independently and it radically differs not only from any natural facts but any other facts (such as supernatural facts) that might have similar features (for more discussions see Shafer-Landau 2003, Cuneo 2007, FitzPatrick 2008, Enoch 2011, Wielenberg 2014, Cuneo and Shafer-Landau 2014, and Skarsaune 2015). Given that non-naturalism conceives normativity as intrinsic and necessary features of moral properties, I suspect that defending an externalist position on such a construal is problematic (I understand a non-naturalist as a born-internalist), because it requires an in-built motivational claim

---

[36]Construing moral realism in this way seems to be attractive because it offers us a metaphysical space (*metaphysischer Spielraum*). Thus, we can plug in any plausible metaphysics of morality that fulfils the defining features from our foregoing discussion. Contrary to the allegation that moral realism binds us to some sort of metaphysical dogmatism, I think that we are free to develop versions of moral realism based on our conception of metaphysics. This attitude, in addition to the fact that our practice of moral thought and discourse is supported by common sense, gives us reason to adopt moral realism insofar as it is grounded on a plausible metaphysics of morality. Brink elegantly expresses this view in the following words: "Moral realism should be our metaethical starting point, and we should give it up only if it does involve unacceptable metaphysical and epistemological commitments" (Brink 1989, p. 24).

(and this is exactly what externalism denies).[37] Thus, I will not pursue this issue directly (my concern is naturalist defence of internalism). Further, while the failure of internalism might not affect the metaphysics, hence the objectivity-claim of naturalism (unless one defends a reductive naturalism), it would affect non-naturalism because its normativity is in-built on the nucleus of moral properties. Alternatively, if non-naturalism fails on the grounds of the mysterious character of its moral properties and facts, then it would be difficult to sustain its motivational claim, hence, its internalist claim.

Moral Naturalism

Naturalism comes in various forms in philosophy; and in some subject-matters, it comes with different dosages of metaphysical commitments. But there is one main challenge that confronts naturalist positions (be it in the philosophy of religion, epistemology, philosophy of mind) namely, how to characterize the term 'natural'. It should not come to us as a surprise that a similar challenge replicates in the moral domain, precisely in metaethics.[38] Although moral naturalism (hereafter simply naturalism) in moral realism maintains that moral properties and facts are natural facts, it is unclear how these 'natural'

---

[37] Nonetheless, Shafer-Landau thinks that there is a possibility of defending motivational externalism within a non-naturalist realist framework (for further reading see *Moral Realism: A Defense*, 2003).

[38]It is also good to note that moral naturalism goes beyond moral realism as it can be defended on the irrealist platform. Jonathan Dancy (2006) observes that, "one can be an ethical naturalist without being a metaphysical naturalist, and vice versa" …for example, "A metaphysical naturalist who thinks that morality is not a matter of fact will neither assert or deny ethical naturalism" (Dancy 2006, p. 122). From both the realist and irrealist perspectives, it does seem that naturalism attacks moral non-naturalism, though for different reasons. Generally, they accuse the non-naturalist of assigning an unnecessary ontology to morality. Simon Blackburn captures this concern thus:

> To ask no more of the world than we already know is there—the ordinary features of things on the basis of which we make decisions about them, like or dislike them, fear them and avoid them, desire them and seek them out. It asks no more than this: a natural world, and patterns of reaction to it. (Blackburn 1984, p. 182)

However, irrealism parts way with naturalist moral realism when it advises us to abandon moral facts. On the contrary, the naturalist moral realism strongly believes that although the non-naturalist project (with its obscure ontology) places morality under threat, we do not have sufficient reason to abandon moral realism entirely. In fact, it is believed that the rejection of the ontological assumptions of non-naturalism makes moral realism a plausible metaethical theory. In this sense, the naturalist moral realist is on a mission to save moral realism.

entities and relations are to be understood. Generally, the *metaphysical claim* of naturalism is construed as follows:

(MN): The objective and mind-independent moral entities (such as moral properties, facts, and relations) are part of the natural world.

(MN) points to the main assumption driving naturalism. Namely, it is believed that moral properties such as goodness, badness, wrongness, rightness, justice, etc., are natural properties, hence part of the natural world.

Reductive Naturalism

However, naturalists differ on the issue of the relationship between natural moral entities and natural non-moral entities. To this end, there are two main positions: on the one hand, some naturalists argue that moral entities are reducible to non-moral entities (Jackson 1998; Railton 1989, 2003a, 2003b; Smith 1994). On the other hand, some contend that although these entities are natural, they are irreducible to non-moral natural entities. The former position is often referred to as *Reductive Naturalism*. Very roughly, it holds that for every moral entity there is a corresponding natural entity. On this view, to locate moral property 'J' means nothing but identifying its corresponding natural equivalent 'D'; and once we substitute J with Justice and D with equal distribution of goods, then Justice equals an equal distribution of goods. Further, to study these natural moral properties and their resultant instantiations we do not need to go beyond the scope of natural sciences.

Nevertheless, this position is not plausible in a number of ways. If moral entities are identical with non-moral entities, it is unclear how the moral entities stand out normatively. This position blurs the distinction that exists in the different construal of natural entities: The construal of naturalness in other natural sciences such as chemistry, biology, physics etc. differs from the naturalness in the moral domain. There is a significant way the natural entities of these other natural scientific fields are different from moral natural entities. The reductive naturalist collapses the ontological distinction between naturalness of moral entities and naturalness of other natural sciences, and in so doing; he mistakenly sees unity where there are differences. Further, it is unclear if the

naturalness of reductive naturalist is actually a natural property. On this ground, Schaber writes:

> Die Frage ist, ob die Tatsachen, auf die sich der Ausdruck 'moralisch gut' bezieht, wirklich natürliche Tatsachen sind. Ich glaube, dass sich auf diese Frage keine befriedigende Antwort finden lässt; und das hat für den reduktiven Naturalismus unangenehme Konsequenzen.[39] (Schaber 1997, p. 103)

To this end, suppose we reduce goodness to 'Glück' – happiness, Schaber asks: Is happiness actually a natural property, the kind that natural sciences study?

> Was heißt 'natürlich'? Wenn wir das als natürlich ansehen, was Gegenstand der Naturwissenschaften sein kann, dann handelt es sich bei dem, was zum Glück von Personen beiträgt, wohl kaum bloß um natürliche Tatsachen. Freundschaft z. B. kann zum Glück von Personen beitragen; und Freundschaft gehört eindeutig nicht zum Gegenstandsbereich der Naturwissenschaften. Es lassen sich hier auch künstlerische Arbeit oder beruflicher Erfolg nennen; das sind Phänomene, die nicht durch die naturwissenschaftliche Forschung erfasst werden. Das zeigt, dass der Naturalist den Begriff der natürlichen Tatsachen (mit denen moralische Tatsachen identisch sein sollen) nicht in einem engen Sinn (oder naturwissenschaftlichen Sinn) verwenden kann.[40] (Schaber 1997, p. 104)

It goes without saying that the reductive naturalist construal of moral entities as identical to natural entities (studied by the natural sciences) triggers off scepticism, not even from the moral domain, but from other natural scientific domains. It is still doubtful if the natural scientists would happily accept that moral entities are studied by natural sciences. In other words, if moral entities were to be identical and reducible to natural entities, the

---

[39] "The question is if the facts which underlie the expression: morally good are actually natural facts. I believe that this claim does not provide a satisfying answer; and this has uncomfortable consequences for reductive naturalism" (Schaber 1997, p. 103 translation is mine).

[40] "What is natural? When we consider natural as the property of natural sciences, then what contributes to persons' happiness is less likely to be construed as natural facts. Friendship, for example, contributes to persons' happiness; but friendship certainly does not belong to the domain of natural sciences. This also applies to artwork or professional success: these are phenomena that are not covered by natural scientific investigation. This shows that the naturalist cannot use the term natural facts (as identical with moral facts) in a strict sense (or in natural scientific sense)" (Schaber 1997, p. 104 translation is mine).

reductive naturalist still saddles the burden of explaining away the source of normativity of moral entities. It has to explain why other non-moral natural entities are not normatively laden. It might be objected that its normativity is embedded in its being moral, that is, in its moral character. But such a claim fails because it assumes exactly what it claims to prove; and in a sense, reductive naturalism would be accused of assigning a mysterious normative feature to moral facts.

Furthermore, reducing moral properties (for example, justice) to natural entities begs the question against other construals of (justice). Suppose, for the sake of argument, that reductive naturalism is true, notice that reducing J to D (where D is an equal distribution of goods and service) precludes S (where S is the view that justice is bringing certain *badly-off* persons to a threshold where they will be better off). Of course, we can extend this case to other moral properties. However, there is a way out of this objection, namely, by multiplying natural entities for all possible construals of J, such that J equals to D, S, …X, Y, Z. However, such a move is very costly for moral realism, because it leads to relativism, precisely metaphysical relativism that has a deep implication for the objectivity claim as well as the substantial moral principles. In addition, the possible construals of J (i.e. D, S …X, Y, Z) are often mutually exclusive; thus, they are competitors about what moral facts *actually* are. Defending the objectivity-claim on such a platform will in advance suffer a knockdown objection. In sum, it goes without saying that taking this move makes the reductive naturalist unfaithful to his Ockham's razor: "*Frustra fit per plura quod potest fieri per pauciora*".[41]

Nonreductive Naturalism

Unlike reductive naturalism, non-reductionist naturalism (simply non-reductionism) holds that although moral entities are natural entities, they are, however, neither identical nor reducible to natural entities. One of the implications of the irreducible character of moral entities is its *sui generis* state (Boyd 1988; 1989; Brink 1984; Sturgeon 1998a, 2006; Miller 1985). Nonreductive naturalists employ the notion of supervenience as one

---

[41] "It is futile to do with more things that which can be done with fewer" (W. Ockham, *Summa Totius Logicae,* i. 12).

of their main strategies of explaining the relationship between moral entities and natural entities. Generally, supervenience is the view that properties or facts of any meaningful domain (D) are part of natural order (N), such that the properties or facts of (D) are fixed by (N)'s properties or facts. Supervenience tries to establish some sort of ontological relations between two classes of properties. If we consider two properties C and K, the supervenience thesis claims that if (C)-property supervenes on (K)-property, it is impossible for there to be a change in (C)-property without having a change in (K)-property.

Given this, it is believed that it is not possible for a change in the moral entities to occur without a corresponding change occurring in the natural entities. Notice that supervenience is here construed as a relation of metaphysical necessity (it can as well be defined in terms of semantic necessity). Metaphysically, consider two moral agents (Mp1) and (Mp2) in a given world (W1). If (Mp1) judges that x is morally wrong; and x supervenes on natural facts (Nf), then it is metaphysically not possible for a change in x to occur without a corresponding change in (Nf); and it also holds for (Mp2) provided his judgement: x is morally wrong supervenes on (Nf). In other words, this holds for all moral judgements with x content in W1. Further, suppose we have, in addition to W1, other worlds W2, W3, W4 …Wn. It is also metaphysically not possible for the moral judgement: 'x is morally wrong' made in other possible worlds to differ from those of (Mp1) and (Mp2), because the x-fact of the judgements (made in the possible worlds) supervenes on a natural fact that is similar in every respect to (Nf). Given its objectivity claim, it would be problematic for the nonreductive naturalist to hold that the natural fact on which x supervenes in W1 differs from that of W2. If he maintains such a position, one might argue that the natural fact on which x supervenes is responsive to particular groups, for example, cultural or moral communities in W1 and W2 respectively. In order to block this objection, nonreductive naturalism has to adopt a more robust metaphysical supervenience where the moral fact that makes 'x is morally wrong' true would supervene on a natural fact applies to all possible worlds.

Nevertheless, supervenience presupposes the possibility of higher-level realities, that is, the entities that supervene on base-entities. For example, consciousness or related

conscious phenomena can only supervene on neural properties if there are such entities as conscious phenomena. The ontological possibility of such higher-level entities is disputable; and this debate replicates itself in the metaethical domain because of the putative and normative characters of moral entities. Precisely, it is asked: Can moral properties supervene on natural properties? Mackie expresses strong doubt over such a relation; for him:

> What is the connection between the natural fact that an action is a piece of deliberate cruelty - say, causing pain just for fun - and the moral fact that it is wrong? It cannot be an entailment, a logical or semantic necessity. Yet it is not merely that the two features occur together. The wrongness must somehow be "consequential' or "supervenient'; it is wrong because it is a piece of deliberate cruelty. But just what in the world is signified by this 'because'? (Mackie 1977, p. 44)

His emphasis is on 'because': He strongly doubts the existence of such facts that are supposed to make 'deliberate cruelty' morally wrong.[42] For him, there are no such objective moral properties and their purported instantiations. On this reading, Mackie's argument seems to hold, because the higher-level entities (in our case moral entities) cannot trigger off change in the base-entities (natural entities); and the absence of such higher-level entities does not affect the natural entities. If it were not so, it would be difficult for Mackie to deny the existence of moral entities (that is, the double-wise necessity relations between the natural and moral entities). However, Mackie's argument against the supervenient possibility of moral entities is ineffective because he embedded his claim on normativity. To his mind, "if there were objective values, they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe", because of their intrinsic normative character (Mackie 1977, p. 37, see also p. 40).

---

[42] It is good to note that depicting "deliberate cruelty" in terms of a *natural fact* is not uncontroversial, because 'cruelty' is a thick moral concept which straddles the divide between descriptive and evaluative features. Characterizing thick concept is debatable, because the moral realist believes that the descriptive and action-guiding features of such concepts form a single moral content; whereas the noncognitivist thinks that the descriptive and action-guiding features of thick concepts come apart.

This objection seems to threaten naturalism, but on a closer look, one discovers that it is a problem for reductive naturalism and internalist nonreductive naturalist. To the former, since he reduced, thereby making moral entities identical to natural entities, it is problematic to explain away whence moral entities acquire their normativity. The reductive naturalist saddles even a heavier burden if he, in addition, holds that moral entities have in-built motivational force. Although he conceives moral entities as irreducible entities, the latter also faces a similar challenge for adopting internalism, the view that it is impossible for moral judgements or belief to occur without corresponding motivation (we shall return to this in chapters 4 and 5). If, however, the nonreductive naturalist rejects internalism, then Mackie's objection would disappear. Suppose we claim that the motivation or in Mackie's term 'to-be-pursuedness' is not built into objective moral facts expressed in moral beliefs (Brink 1986, 1989; Schaber 1997; Zangwill 2003 defended such view, and it is the position defended in this present work), then there would be nothing queer for moral entities to supervene on natural entities.[43]

3.5. Conclusion

Thus far we have outlined four features of moral realism:

(I)   Moral claims (such as judgements and statements) are capable of being true or false.

(II)   Moral claims are truth-apt because there are objective moral facts to which they refer.

---

[43] Brink argues that: "There is nothing strange or certainly nothing unique about the supervenience of moral properties upon physical properties. Assuming materialism is true, mental states supervene on physical states, yet few think that mental states are metaphysically queer (and those who do not think that supervenience makes them queer). Social facts such as unemployment, inflation, and exploitation supervene upon physical facts, yet no one supposes that social facts are metaphysically queer. Biological states such as being an organism supervene on physical states, yet no one supposes that organisms are queer entities. Macro-scopic material objects such as tables supervene on micro-scopic physical particles, yet no one supposes that tables are queer entities. In short, it is difficult to see how the realist's use of supervenience in explaining the relationship between moral and physical properties makes this position queer" (Brink 1984, p. 120, see also Brink 1989, pp 177 - 180).

(III)   These moral facts exist (and continue to exist) mind-independently.

(IV)   At least some of the moral claims are literally true because they depict these facts correctly.

For the purposes of our work, we take (I) to (IV) as the defining features of moral realism. However, we shall construe as moral facts in (III) as natural entities, which are neither reducible nor identical to natural entities. Based on this, in the second part of this project, we attempt to answer the following question: 'In what sense, if at all, can objective moral judgement (that is, understood in the moral realist terms) motivate agents to act morally?' In chapter one, we argued, contrary to Hume's charge, that moral belief motivates; if it were not so, moral realism will be morally impractical, hence motivationally impotent. However, the issue at hand is how to reconcile the connection between moral judgement and motivation within the moral realist framework. I consider the realist internalist solution, namely, that moral judgement necessarily and efficaciously motivate us. Essentially, this position presents an incomplete picture of human psychology. I argue that the realist motivational internalist saddles the burden of explaining away the role of desires in moral actions in particular, and human psychology in general. To fill this gap, I expound and defend externalism. I argue that *moral care*, *moral-self* and *moral desires* are essential to moral motivation. Moral motivation does not stem directly from moral judgements; rather from an independent desire, namely, a moral desire for realizing the moral values we care about, that is, the desire to be faithful to one's moral-self.

PART TWO

Psychology of Moral Motivation

Chapter 4

Motivational Internalism

> *. . . [O]ne can be indifferent to morality without error.*
>
> **Philippa Foot, 1978**

4.0 Introduction

In this chapter and the next one, I argue that the claims that [1] moral judgements are motivationally efficacious, and [2] the connection between moral judgement and motivation is necessary, are false. In other words, satisfactory solutions to the moral motivational problem must be found for moral realism; and externalism can provide them. However, before we embark on this journey, a stage-setting is necessary. The debate between *internalism* and *externalism* has taken various turns in recent years; and because of lack clarity of concepts, the debate has led to more confusion (Brink 1989, p. 40; Cohon 1993, p. 266; Darwall 1992, p. 155; Mele 1996, p. 728; Tresan 2009, p. 52). In some of the cases, even the internalists seem to be talking past one another partly because of various characterizations of internalism. Internalism is the view that moral claims are both internal and necessary. This general characterization is not very helpful because internalism comes in various flavours. Hence, for the purposes of this work it is in order to distinguish the various internalist positions. First, internalism is a position that focuses on the justifying *reasons* or *motivating reasons* for action that are internal to the agent. In the former, the argument is that what counts for or against actions is internally embedded on the agent's subjective psychological profile. Bernard Williams, for example, argues that agents' reasons for action are always *internal* insofar as such reasons are rooted in the subjective motivational set or desires (Williams 1981, pp. 102, 109). On this view, necessarily an agent cannot act if she does not have a reason. In other words, in the presence of reasons for action, it is necessarily the case that agents act accordingly (Rosati 2014). This position is often called *reason internalism* or *internalism about practical*

reason.[44] The defenders of the opposing position – reason externalism – think otherwise.[45] In the latter, the main concern is not just about justifying, but also about motivating reasons for action. Unlike reason internalism, the connection here is between moral considerations and motivation. It is generally agreed that moral considerations make motivational claims on us. However, the motivational internalists strongly claim that the motivational force in question is internal to moral considerations. That is, it is necessarily the case that moral considerations motivate agents (Plato; Nagel 1970; McNaughton 1988; McDowell 1979; Dancy 1993; Bromwich 2013). On this level, although reason internalism and motivational internalism are distinct positions, however, it would be false to claim that they do not overlap: "It is true that one set of issues is sometimes appealed to in arguments concerning the other. But issues about motivation and issues about practical reason are *prima facie* distinct" (Zangwill 2008, p. 93).[46] I will be mainly concerned, for the purposes of this work, with the latter – namely that moral judgement internally motivates.

Second, internalism is defended on the bases of the psychological states underlying the moral considerations in question. On this level, motivational internalism is explained either in terms of conative states or cognitive states. In the former, it is the thesis that non-cognitive states such as pro-attitudes motivate agents. Precisely, non-cognitivism strongly claims that these conative states are motivational states. Given that

---

[44] For more discussions on argument for internal reason see Goldman, 2005; Manne, 2014.

[45] For arguments against reason internalism see, for example, Lubin and Dean, 2009; McDowell, 1995; Mason, 2006.

[46] Roughly, it is possible for moral reason (that is, reason for action) to figure in agent's moral judgement and motivation. Philip Pettit and Michael Smith argue that "this distinction between motivating and normative reasons needs to be handled with some care, however…. when certain motivating reasons explain what an agent does they do so by allowing us to see a minimally rational pattern in her conduct. It thus follows that motivating reasons cannot be divorced entirely from considerations of rationality" (Pettit and Smith 2006, p. 7). Nevertheless, they further write, "there is, we think, a quite decisive consideration in favour of making a sharp distinction between the two kinds of reason. An agent may have a motivating reason to φ without having any normative reason at all to φ — indeed, she might think, quite correctly, that all the normative reasons that there are tell against her φ-ing — and she may also have a normative reason to φ without having any motivating reason to φ. Any account of the nature of motivating and normative reasons must therefore preserve at least these two stark possibilities" (Pettit and Smith 2006, p. 7; also for more discussion on the difference between reason for action and reason for motivation, see Alfred R. Mele 1996, p. 731).

moral considerations are expressions of these states, to judge means to express these states, which in turn means to be necessarily motivated. This position is compatible with Humean understanding of psychology of motivation. On this view, motivation is guaranteed by desires. However, it is highly contended whether Hume's motivational claim entails non-cognitivism (Pigden 2009; Joyce 2010). Some argue that he defended an internalist motivational position (Coleman 1992; Stroud 1977; Mackie 1980; Darwall 1983); whereas others maintain that Hume is an externalist (Brown 1988). In the latter, motivational internalism is the view that motivation is internal to moral considerations (where moral considerations are understood as cognitively-laden). Defenders of this position appeal to cognitive states, for example, reason and its product beliefs, to explain the necessary connection between moral considerations and motivation. On this view, cognitive states motivate not only necessarily, but sufficiently as well (Korsgaard 1996; Nagel 1970; Scanlon 1998; Platts 1980; Bromwich 2009, 2013).

The philosophers of the latter group espouse moral cognitivism (in addition, most of them are moral realists). Most of them have a common goal, namely to reject Humean psychology of motivation. According to them, the claims that reason is inert, and that moral beliefs are not sufficient (because they require pre-existing desires) to motivate agents are false. Nevertheless, moral cognitivists as well as moral realists are sharply divided on the issue. Some philosophers such as Brink (1989), Zangwill (2003; 2008) and Shafer-Landau (1998) argue that we need not reject Humeanism to explain moral motivation. To their mind, moral cognitivism as well as moral realism is compatible with Humean psychology of motivation.[47] Following this position, I offer a solution to the moral motivational problem, which holds that distinct moral desires provide the motives for acting on moral judgements (chapter 7). Among other things, the position I defend in this work does not relegate the moral agent as a moral person to the background. It takes into account the various degrees of moral beliefs and the strengths of desires, depth of traits, etc., figuring in the moral agent's motivational profile. However, given that motivational internalism, on the cognitivist level, is characterized differently (even in

---

[47] Some philosophers such as Smith (1994) hold both realist internalism and Humean theory of motivation. As we shall see in chapter 4, he argues that the realists who are against Humeanism are, in fact, wrong.

some cases, these construals overlap, thereby blurring the significant claims of motivational internalism), I will, in the first section of this chapter, walk through a number of these various understandings. This important exercise will enable us to identify the underlying features of motivational internalism, and thus, help us to characterise the versions of realist internalism relevant for the purposes of this work.

4.1 Motivation Internalism

One of the features driving the motivational internalist position is the idea of necessity.[48] The *Necessity Claim* of motivational internalism (hereafter, simply MI) is the view that the connection between moral considerations and motivation is unconditional. Motivation does not follow from moral considerations contingently. As we shall see later, the necessity claim comes in metaphysical or conceptual form. Generally, moral considerations have different contents such as moral properties, moral facts, moral judgements, moral beliefs, etc. Suppose we take moral properties as an instance of moral consideration. Hence, on this construal, the motivational internalist argues that the knowledge of moral properties necessarily guarantees motivation.[49] On this assumption of MI, Plato once held that *knowing* the good is necessarily *doing* the good. We can formulate this version of MI as follows:

> *Platonic Internalism:* Necessarily, if an agent *knows* that φ is morally good, then she is moved to act accordingly.

Plato builds his MI on the notion of knowledge of moral goodness – moral property. Mackie considers Plato's construal of moral properties as the instantiation of objective moral properties underlying moral realism. According to him,

---

[48] These features are not peculiar to internalist moral realism; they also apply to non-cognitivist version of internalism. However, we shall consider them here within the moral realist internalist construal.

[49] We can, if only roughly, classify this position under Darwall's non-constitutive existence internalism (Darwall 1992). As we shall shortly see, unlike the constitutive existence internalism, the non-constitutive version argues that motivation necessarily guaranteed through epistemic contact.

Plato's Forms give a dramatic picture of what objective values *would have to be*. The Form of the Good is such that knowledge of it provides the knower with both a direction and an overriding motive; something's being good both tells the person who knows this to pursue it and makes him pursue it. (Mackie 1977, p. 37)

As we observed in the last chapter, one cannot attribute this position to moral realism in general, because such 'objective' facts or values do not fit into the frameworks of various forms of moral realism, for example, ethical naturalism. These facts apply to moral realism insofar as one accepts an *inflationary ontology*. That is, the view that moral properties such as goodness, are neither visible to the senses nor detectable by the natural sciences; hence, they are non-natural moral properties (Putnam 2004, pp. 17 – 19). In other words, Mackie's charge is only true for the non-naturalist saddling the platonic metaphysical hangover. Therefore, Mackie's construal of objective moral facts is false, because it reduces moral realism to Platonic Realism.

However, Platonic realism is disturbing at least on two main grounds. First, ontologically, it assigns a rather strange character to moral facts and properties. Indeed, Mackie was right when he said, "If there were objective values then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe" (Mackie 1977, p. 38). Second, epistemologically, it reduces the accessibility of moral facts and properties to special intuitive faculty. Again, Mackie says, "Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else" (Mackie 1977, p. 38).

It is on these *moral ontological* and *epistemological grounds* that Plato built his internalist position. Moreover, on Mackie's reading, Plato's moral goodness is not only action-guiding, but it as well provides moral agents with the motivating reasons for action. Notice that motivation is tied overridingly to moral goodness, such that if an agent knows that something is morally good, he is necessarily motivated to do it. The *moral whyness*, namely reasons for acting is embedded on the moral properties. Hence, the *reason-giving* and *motive-providing* characters of moral properties rule out the possibility of overridingness by other competing desires, motives or reasons. Bromwich describes

this form of MI as *decisive internalism* (Bromwich 2009). Nevertheless, Plato is not alone on this podium; at least the early rational intuitionists espouse the idea of motivation based on some sort of perceptual and epistemic contacts.[50] Price, for example, argues thus:

> All men continually feel, that the perception of right and wrong excites to action. It seems extremely evident, that excitement belongs to the very ideas of moral right and wrong, and is essentially inseparable from the apprehension of them. […] When we are conscious that an action is *fit* to be done, or that it *ought* to be done, it is not conceivable that we can remain *uninfluenced*, or want a *motive* to action. (Price 1965, Ch. VIII § 706)

To Price's mind, moral action is tied experientially to our perception of moral rightness or wrongness. In other words, it is a part of our moral phenomenology that excitement or motivation is tied to the perception of moral rightness or wrongness. Notice that his emphasis is on the universality – *all men*; and persistence – *continually feel*: both elements point to the necessary connection between moral considerations and motivation (we shall return to these elements shortly in 4.3). In other words, it is the idea of the inseparability of motivation from moral rightness and wrongness that makes it inconceivable for an agent to perceive a moral ought and yet remain unmoved.

First, the problem with Platonic Internalism is that it requires that agents possess special intellectual faculties to perceive moral facts and properties. Hence, agents lacking such faculties would be generally motivationally unfit. Moreover, it is understood that such epistemological demands restrict the accessibility of moral properties. Second, although it guarantees motivation, that is, it explains the connection between moral considerations and motivation; it fails to account for motivational failures, which are as

---

[50] We can classify the early British moral intuitionists (such as Samuel Clarke, John Balguy, and Richard Price) under this group. Darwall writes, "They affirm the ancient Platonic doctrine that it is of the nature of the ethical that it cannot be known or perceived with indifference. And they explicitly deny that what makes a moral claim true, even a claim about obligation, is anything essentially to do with motivation or the will" (Darwall 1995, p. 111).

well part of our moral experience. In other words, it is unable to justify some cases (e.g. weakness of will, overridingness of stronger emotions) which are significant parts of our moral experience.[51] Notice that Price tied motivation to perception of right and wrong. However, since perception does not necessarily require cognitive activities such as belief formation or linguistic apparatus, they might be cases where moral perception precludes formation of moral beliefs.[52] That is, in a way that does not involve an agent's formation of first-person moral beliefs. This notion of MI does not fit the purposes of this work, because it says, "not that ethical belief or sincere assertion necessarily motivates, but that actual consciousness of or cognitive contact with the ethical does" (Darwall 1992, p. 157).

More so, apart from basing motivation on moral properties, it can as well be tied to moral obligations. Generally, moral obligations contain facts about what is to be done; but such facts alone do not necessarily engage agents in a morally significant way.[53] For example, it is a fact that torture is morally bad, but such mere facts might not be appealing to agents (in particular cases) in a way that makes the content of the obligation motivationally significant.[54] W.D. Falk makes this point clear: He argues that moral obligations normally have *grounds,* but such grounds are not the *kernel* of the moral obligation.

---

[51] McDowell's MI appears to suffer a similar threat. According to him, the virtuous person sees or perceives moral situation in a special way; such that he is necessarily motivated by such perception (1978, 1979, and 1981). The virtuous person, on McDowellian analysis, is overridingly or decisively motivated. Given that, he tied virtues to cognitive capacities; it goes without saying that the agent's motives might clash or even be toppled by stronger competing desires, motives or other conative states.

[52] Moral perception is, after all, a species of perception, not something radically different.

[53] Mele argues that agents might be morally obliged or required to do something without being motivated because the content of such moral obligation is not necessarily motivationally efficacious. "This view is problematic for moral cognitivists. Presumably, on a cognitivist view, Hitler may have been morally required not to order the extermination of his victims even if Hitler never had a motivation, in the identified sense, not to order their extermination" (Mele 1996, p. 729).

[54] As Zangwill reads him, Dancy seems to be arguing that bare moral facts, independent of underlying psychological states, motivate agents. "He says that one is usually moved by facts in the world, not by psychological states (Dancy, 2000). I suspect that this amounts to the thought that our motivational states usually have non-psychological propositional contents" (Zangwill 2008, p. 93). If this view were to make sense, then Dancy has to, at least provide the role of agents' psychological profile in moral motivation, precisely how such states figure in moral judgements. In other words, he has to explain away the possibility of, say, two agents with the same moral factual content, but having varying, even if minimal, dispositional or psychological states.

> The nature of the things, which we are obliged to do contains only the grounds, but not, yet the *essence* of moral obligation. What alone can render a prospective action obligatory is that an agent is in some manner impelled to do it, or that he thinks he would be so impelled if he reflected. [...] Hence what makes the good act a duty is not the bare fact that it would be good when done, but the fact that the thought of it is related to ourselves in a special manner; and even if it were the case that ultimately none but good acts were obligatory, their goodness would be no more than the *ground* of a separate obligation to do them. (Falk 1945, p. 145)

Notice that this view departs from Platonic view, at least on two grounds: First, it does not claim that motivation is closely connected to perceptual or epistemic contacts of moral rightness and wrongness. Second, it argues that moral facts and properties are separable from moral obligations and duties. The former are bare facts that are external to agents[55] whereas the latter internally engage the agents' psychology. According to Falk, it is in the nature of moral obligation to internally impel agents to act in accordance to what is required to be done. Hence, he embeds moral motivation on moral duties or obligations. Furthermore, he likens agents under a moral obligation to persons "having an impulse" or "having an obsession" (Falk 1945, p. 143). For example, if an agent is obliged not to torture an immigrant, on this construal, she or he is internally impelled to act accordingly: "to be obliged is to affect, to be obliged is to be affected" (Falk 1945, p. 143). To illustrate this form of MI,

> *Moral Obligational Internalism:* Necessarily, if an agent is obliged to φ, she is affected or influenced, hence motivated to φ.

The agent's impulse is grounded on the reasons provided by moral obligations. However, it is good to note that Falk did not make this distinction (Darwall 1992, p. 155). In any case, the obligatoriness of what is to be done motivates agents. On this view, the necessity

---

[55] This implies that Falk is an externalist about moral facts and properties; but an internalist about moral obligations and duties (see Mahon 2005).

claim of this form of MI is tied to the inseparability of motives from moral obligations. This view is akin to the Kantian idea of the inescapability of moral duty (Timmons and Baiasu 2013, p. 150; see also Falk 1986, p.29). Hence, on this construal, it will be impossible for someone to accept obligatoriness – the essence of moral obligation – without being impelled, affected or influenced, that is, motivated to act accordingly.

The problem with this version of MI includes the following: first, consider a community espousing a number of moral obligations. If we assume, given our corresponding experience in real life, it is at least possible that some of them will not be motivated to X. This is because the motives that moral obligations provide might differ from the agents' actual motives; and in such cases, the obligatoriness of what is to be done might fail to motivate efficaciously. Hence, this version of MI fails to capture the difference between the essence of moral obligations and agent's actual or hypothetical motives.

Second, Falk claimed that moral obligation has obsessional force, such that agents who are under moral obligation are impelled to act accordingly. While it is true that this construal of MI explains the connection between moral obligation and motivation, it fails (due to the obsessional character of moral obligations) to make provisions for the members of the community who might have general motivational problems. For example, there might be members of the community who genuinely understand (and even accept) what is morally required to be done; but it is also possible (actually and hypothetically) that they may lack the basic motivational ingredients in their psychological profiles as in the cases of pathology, "mental block or mental defect, or a distraction" (Cohon 1988, p. 112). Finally, for the purposes of this work, we are interested in forms of MI that are both agentic and judgement-based.

4.2 From Judgement to Judgement Internalism

Before we proceed, some remarks are in order. What does it mean to judge morally (and even otherwise)? This question is debatable in metaethics. The difficulty in defining

moral judgement lies on how its content is to be characterised. For example, when an agent 'A' judges x as morally wrong, how is her judgement to be understood?

On the moral realist construal, the state underlying A's judgement of x is taken to be cognitive; and as we saw in the previous chapter, the noncognitivists think otherwise. For the purposes of this work, we shall restrict our concern to moral realist understanding of moral judgement. Here judgement is on par with the ordinary act of judging; hence, by judging something to be morally wrong, an agent is endorsing or affirming a state of affair in the world. For example, by judging that torture is morally wrong, 'A' is both asserting and affirming something about the wrongness of torture. On this view, such affirmation entails belief: 'A' affirms that torture is morally wrong because he *believes* that the act is morally significant. 'A's judgement expresses a belief about x; and this belief is, in turn, cognitive-laden. To this end, the realist internalist argues that such belief-based moral judgements are motivationally efficacious. That is, they motivate agents independent of antecedent or mediating desires or other non-cognitive states. Hence, the main objective of this chapter and the next one is to argue that belief-based moral judgements (hereafter, simply moral beliefs or judgements) do not motivate alone. Precisely, I argue that the realist motivational internalist solutions on both robust and conditionalized construals fail, because motivation is not internal to moral judgements. In this chapter, I will consider the plausibility of a robust construal of internalist moral realism. Generally, we can now formulate this view of MI as follows:

> *Realist Motivational Internalism (RMI)*: Necessarily, if agents *judge* or *believe* that they are morally required (or morally ought) to φ, they are motivated to φ.

On this construal, motivation is based on agents' judgement engagement; and therefore, it is no longer mere knowledge of moral properties or perception of right and wrong that gives agents motives. Darwall rightly observes that this form of MI has a space for agents' engagement in "deliberative process of practical reasoning and judgement" (Darwall 1992, p. 158). More importantly, however, MI on this analysis is constitutive: It is argued that moral judgements constitute motivation. The content of moral judgement is capable of providing motivation without relying on independent desires. For example, it is

strongly argued that moral judgement does not require a combination of non-moral contents (not even prudential content) in order to motivate agents. In other words, the content of a moral judgement is motivationally efficacious on its own (Bromwich 2009; Mele 1996; Darwall 1992). Brink refers to RMI as *appraiser internalism*. That is, the claim that

> It is in virtue of the concept of morality that moral belief or moral judgement provides the appraiser with motivation or reasons for action. Thus, it is a conceptual truth about morality, according to appraiser internalism, that someone who holds a moral belief or makes a moral judgement is motivated to, or has reason to, perform the action judged favourably (Brink 1989, p. 40).

The content of moral belief makes theoretical claims as well as motivational claims on the agents. Hence, RMI is to be understood as the thesis that moral belief guarantees motivation insofar as agents hold the content of such belief as true. The guaranteed motivation "rests upon the nature of belief itself and upon the content of the belief that one is (oneself) morally required to…." (Mele 1996, p. 729) In other words, it is by believing that something is morally bad (or morally good) that agents are said to be *judges* or *appraisers*. The realist internalist, in turn, believes that such moral judgement possesses motivational force to move agents accordingly.

Furthermore, the necessity claim of MI carries some sort of metaphysical commitment.[56] If only roughly, this claim is supposed to apply to *all persons* and *possible worlds* sharing the concept of morality. "To get internalism we must posit accompaniment, not just actually, but throughout possible worlds. That is, 'Entail' indicates that the accompaniment is necessary" (Tresan 2009, p. 54). Recall RMI, and when applied to this case, it would imply that *all* agents judging or believing that φ is morally required would be motivated to φ (at least if φ is understood as *normatively unqualified)*. The claim amount to this, it is necessary that any agent in any possible world who judges or believes that φ is morally required is motivated to φ. This claim holds

---

[56] Once again, the connection here is understood from the moral realist perspective.

provided the agents' content of belief is the same (or at least similar) across possible worlds.

However, notice that it does not say anything about the agents' psychological profiles. Suppose we characterise their psychological profile as "normal", then, on the internalist analysis, the content of moral belief coupled with the act of believing would motivate globally. Hence, according to internalism, it is impossible for a change to occur in the content of an agent's belief without a corresponding change in her motivation. To summarise, granting the necessity claim, the internalist argues that necessarily if any agent believes or judges φ to be morally required, she is motivated to φ. Moral belief on its own has motivational efficacy. It generates and guarantees motivation independent of any other desire or non-moral content. I refer to this claim as the core claim of moral realists' MI. Moreover, if they are, in turn, correct then internalists would not just be "advancing a reformative conception of cognitivist belief and alternative to a Humean theory of motivation" (Mele 1996, p. 736), but also their move would constitute a knock-down objection to Hume's picture of human psychology of action. However, we shall see shortly that this view of moral realist MI is false. Before then recall Hume's claim:

(1) (Moral) beliefs do not motivate.

For the moral realist internalist to prove that Hume's constraint on belief is false, he has to justify how moral beliefs motivate. However, his justification has to be at least on the same strength of attack levelled against it. Notice that Hume's attack is not just that moral beliefs do not motivate, but that genuine beliefs do not motivate all. Therefore, the internalist rebuttal must not be that moral beliefs or that some of them *can* motivate. But they *must* motivate (Shafer-Landau 2000, p. 279). Danielle Bromwich, for one, argues that realist internalists have reasons to charge against Hume's constraint without diluting their position. According to her, all beliefs, not just moral beliefs, motivate *simpliciter*.[57]

---

[57] She puts this thus: "In defence of this thesis (that is, cognitivist internalism) it is tempting to either argue that the Humean constraint only applies to non-normative beliefs or that moral beliefs only motivate *ceteris paribus*. But succumbing to the first temptation places one under a burden to justify what is motivationally exceptional about moral beliefs and succumbing to the second temptation saddles one with a thesis that

Since the success of anti-Humeanism depends on refuting (1) without admitting defeasibility (Bromwich 2009, 2013), I will in this chapter focus on the plausibility of the motivational efficacy of moral beliefs, and the necessity claim of realist internalism.

Furthermore, if Bromwich's claim is correct, then moral belief does not just motivate simply because it is moral, but because it is essentially belief. Notice that this position implies that:

(2) All beliefs, not just moral beliefs, motivate.

It also claims that moral beliefs motivate indefeasibly, hence

(3) The content of first-person moral judgement guarantees motivation indefeasibly. [As we shall see later, robust internalism differs from conditionalized internalism concerning the scope of indefeasibility claim].

Assuming that (2) and (3) are true, then

RMI*: It is necessary that, for any agent A, and for any action φ, if A judges that she is morally required (or that it is right) to φ, then she is motivated to φ.

In other words, the realist internalist can only show that (1) is false by proving that (2) is true not on per *ceteris paribus* basis. I will refer to this position as *Robust Realist Internalism* as opposed to *Conditionalized Realist Internalism*. Suffice for now to say that being an anti-Humean per *ceteris paribus* not only begs a question against Hume's constraint in (1), but it also carries the cost of justification on a number of significant ways as well (Miller 2003; Strandberg 2013; Bromwich 2009, 2013). In this chapter, I will consider internalism in its standard construal;[58] whereas in the next chapter I will

---

fails to do justice to the practicality intuition that cognitivist motivational internalism is supposed to capture" (Bromwich 2009, p. 2).

[58] Standardly, internalism is understood as a *per simpliciter* claim, that is, the connection between moral judgement and motivation is internal and not defeasible. Brink elegantly supports this claim thus:

> Though it is unlikely that any interpretation of internalism [...] will be faithful to everything every party to debates about moral motivation has claimed, our interpretations should be guided at least in part by what central figures in the debate have claimed, including what they have claimed about the mutual relations among cognitivism, internalism and desire-base views of motivation. First, I think that many parties to debates about moral motivation have in fact accepted and relied on my stronger formulation of internalism16 [non-defeasible motivational internalism]. This certainly

consider the conditionalized realist internalist solution. I argue that internalist solutions fail on both fronts. However, before we take on Robust Internalism (hereafter simply RI), it is important to consider two more core features of MI.

4.3 Internality and Conceptuality

Apart from the necessity claim, there are, at least, two other main features relevant to understanding the claims of MI, viz., *Internality* and *Conceptuality*. First, the internality claims: It is closely tied to the necessity claim. Recall that the internalist believes that the connection is necessary, such that motivation unconditionally, and not contingently, accompanies moral considerations. On the construal of realist internalist, motivation is internal to moral judgement. Mele writes that "…what is guaranteed, more precisely, is that motivation […] is built into any belief that one is (oneself) morally required to [...] and is internal to the belief of that kind in this sense" (Mele 1996, p. 730). It is this in-built force that explains why any agent is necessarily motivated upon believing that he is morally required to do something. In *Obligation and Motivation in Recent Moral Philosophy*, W. K. Frankena points to this internality claim as follows: It is so rooted in moral considerations that it is logically impossible for agents not to be motivated even if they lack actual or dispositional motives for doing what is morally required (Frankena 1958, pp. 40 – 41). Suffice to say that unlike the realist, the non-cognitivist locates the internal or built-in motivational force of moral judgement on some sort of complex psychological states (namely, underlying desires, feelings). Notice that in order to justify (2) the realist internalist has to root such intrinsic or built-in motivation in moral beliefs alone. Zangwill was right when he argued that "the internalist needs to claim not just that moral beliefs are necessarily motivating, but that motivation is essential to moral beliefs" (Zangwill 2008, p. 94). Although it is not our concern here, but it is good to briefly note that some cognitivists argue that motivation is located in a sort of hermaphrodite

---

seems to be true of the noncognitivists. Second, only the stronger version of internalism makes the puzzle genuinely inconsistent; because Smith's formulation of the problem employs a weaker version of internalism, his triad is not genuinely inconsistent. (Brink 1997, p. 8; see also Tresan's characterization of judgement internalism in *Metaethical Internalism: Another Neglected Distinction*, 2009, p. 53 - 54)

psychological states called *Besire* (Zangwill ibid., p. 51).[59] Another way of making the internality claim of MI is to ask whether motivation is embedded on the content of moral belief or not: In response to this Roskies writes that motivation "must hold in virtue of the content of the moral belief itself, not in virtue of some contingent or auxiliary non-moral fact or reason" (Roskies 2003, p. 52). In sum, the necessity claim makes sense only if motivation is in-built, that is, if it is essential to moral judgements.[60]

Second, the conceptuality claim: Internalism has been largely understood as a conceptual claim. That is, the necessary connection is understood as an essential part of our ordinary language and meaning of moral terms (Strandberg and Björklund 2013). If only roughly, just as it is part of the ordinary meaning of terms, for example, to understand a bachelor as an unmarried man, the internalist sees the concept of moral judgement as motivationally efficacious. For example, Nagel argues that

> Motivation must be tied to the truth, or meaning, of ethical statements that when in a particular case someone is (or perhaps merely believes that he is) morally required to do something, it follows that he has a motivation for doing it. (Nagel 1970, p. 7)

The conceptual claim seems to reflect the folk intuition about moral motivation. Later in this work (chapter 6), I will argue that the internalist claim, namely, that folk intuition is on the side of internalism, is false. Precisely, I argue that folk intuition about moral motivation is inconclusive.

---

[59] This position flies in the face of ontological distinctions. The defenders of Besire think that it is possible to collapse the ontological wall between belief and desire. Zangwill writes that "Besires are not supposed to be gerrymandered mereological sums of two states – a belief plus a desire – each of which could occur without the other. Besires are conceived of as unitary states that have both the representational characteristics of beliefs and the motivational characteristics of desires" (Zangwill 2008, p. 51). Suffice to say that besire-project not only blurs the ontological uniqueness of beliefs and desires, but it as well muddles the distinctive causal roles of psychological states in moral practice.

[60] According to Fine (1994), the necessity claim does not entail the essentiality claim, because it is possible for the former to hold without the latter. For example, an internalist can believe that there is necessary connection between moral judgement and motivation, while denying that such motivation springs from moral beliefs. Internalists are divided on this issue: The internalists of the anti-Humean tradition are more likely to hold that motivation is essential to moral beliefs alone; whereas the internalists of the Humean tradition believe that desires are required for motivation, where the desires are intrinsically generated by moral beliefs. In any case, following Zangwill, I argue that MI should be construed not only in terms of the necessity claim, but also as an essentiality claim because "if motivation is essential to moral beliefs, that would explain why moral beliefs are necessarily motivating" (Zangwill 2008, p. 95).

4.4 Robust Realist Motivational Internalism

The version of internalism relevant to this chapter does not understand the motivational force of moral belief as indefeasible in terms of overridingness. The motivation guaranteed does not mute other competing motivations. It does not also claim that moral judgements furnish agents with overriding motivation, which in turn leads to action. Rather, it holds that the *necessary* connection between moral belief and motivation is *not defeasible*. To illustrate,

> RMI**: It is necessary that, for any agent A, and for any action φ, if A judges that she is morally required (or that it is right) to φ, then she is *efficaciously* motivated to φ by her moral judgement *alone* and not by external desires, feelings or emotions.

It is good to note that the emphasis is on *necessity* and *internality*, hence robust connection: The claim that moral belief motivates by itself, namely that it has motivational efficacy independent of any desires, emotions or inclinations. Although motivation might fail in the face of competing factors, the necessary connection is not defeasible, because motivation is essential, hence in-built in moral judgement. What such robust internality claim tends to block is the problem of creeping (actual or conceptual) external factors – for example, desires, non-moral motives, etc. (Bromwich 2009, p.19). For example, in the case of desires, the robust internalist insists that moral beliefs motivate without depending on any antecedent or mediating desires. This concern is to avoid the temptation of giving in to Humean belief-desire seduction. Hence, in order to secure the necessary connection, the task before the internalist is that of showing how motivation internally lies within moral beliefs. If this strategy were to work, then it would be correct to say that:

> Internalist moral realism [is] a sort of Holy Grail of meta-ethics. It offers us all we ever wanted from morality. The internalist claim gives morality the psychological "oomph" it needs to motivate action by itself, rather than having to hitch [a] motivational ride on pre-or non-moral motives. The realist thesis makes morality what it seems to be: a discourse about facts—moral facts—which we can

discover, about which we can disagree, and of which we can often convince each other. (Noggle 1997, p. 88)

The argumentative strategy of the internalist must be effective in explaining how moral motivation stems from the *content* of the judge's moral beliefs alone. Having rightly accused conditionalized realist internalists of explaining agents' moral motivation based on some sort of "non-moral tendency towards a coherent or actual psychology"[61] (Bromwich 2009, p. 19), the internalists must provide a sufficient picture of how motivation is *solely* derived from moral beliefs. In other words, they have to justify how motivation is internal to moral beliefs without relying on any actual or hypothetical psychology that is external to the content of the moral judgements. On the contrary, in what follows, I will argue that the RMI** is false. Our argumentative strategy is to show that an agent can make genuine first-person moral judgements and yet fail to be motivated. In other words, I argue that the necessity claim embedded on the thesis: Motivation is internal and essential to moral judgement is false.

4.5 Olivia's Case: An Argument against Robust Internalism

Recall Olivia, who believes that environmental harm is morally unjust, and yet declines to sign a petition against the university's unethical environmental policy. Suppose Emma approaches her again – two days after the first incident – appealing to her to sign the petition. This time, Emma substantiates her moral conclusion with concrete evidence. After she finished, Olivia tells her that she never doubted whether the university's policy

---

[61]Bromwich argues as follows:

> Smith points out that, if one lacks a tendency toward a coherent psychology, then one is irrational. Nevertheless, from the cognitivist motivational internalist's point of view, whether or not the subject is irrational is beside the point. The problem, from the realist motivational internalist's point of view, is that without this non-moral tendency towards a coherent psychology moral cognition is motivationally inert. It is not, after all, the subject's understanding of the content of her first person moral judgement that causes the desire to act in accordance with that moral judgement; it is, rather, the freestanding non-moral tendency towards a coherent or rational psychology that causes the desire to act in accordance with that moral judgement. It is difficult to see, then, how Smith really does preserve the idea that motivation must be built into the cognitive moral judgement. (Bromwich 2009, pp. 18 – 19)

was an instance of environmental injustice. She continues to believe that environmental harm is morally unjust and that policy in question is morally impermissible. However, even at that, Olivia does not seem to care about the issue at stake. In other words, she is indifferent about signing the petition (she is indifferent about the moral issue in question).

The phenomenon of indifference is part of human experience. We witness cases where people remain indifferent to various issues, ranging from simple to complex everyday issues. It is not rare to encounter people who do not *care* about what they believe. Indifference is, as well, an essential part of our moral experience. It is not queer to claim that people exhibit indifference in the face of moral demands or issues. Even though it is a contestable position, moral indifference is defended in philosophy (Foot 1972; Stocker 1979; Milo 1981; Brink 1989; Mele 1996; Svavarsdóttir 1999; Zangwill 2008).[62] Zangwill writes: "it certainly seems that moral indifference is no mere abstract philosopher's possibility, but a common actual phenomenon" (Zangwill 2008, p. 102).

The idea of indifference, if only roughly, is about the degree of people's interest or care about what they believe. "Intuitively, we want things more than others, and we believe some things to a greater degree than others. (We are more confident of some claims than others.) Our mental world is not black and white." (Zangwill 2008, p. 95) This experience replicates in the domain of morality, we care about things more than others as well as believe things in different degrees. And the degree with which agents care about moral demands or issues determines to a large extent their motivation. Hence, our strategy is to show that internalism fails to capture this phenomenon in its psychology of moral motivation. Alternatively, if it is the case that the argument from indifference succeeds, then the claim that motivation is internal to moral belief is false given the possibility of holding a genuine moral belief and yet not caring about morality.

Moral indifference is the belief that it is, in fact, possible for someone to *know* or even *believe* that he or she is morally required to do something and yet not *care* about it. The phenomenon of indifference differs in its various construal of caring about the requirements of morality. For example, an agent might be presented as either 'not caring

---

[62] Although Frankena defends an internalist position, he still believes that moral indifference is possible. He writes thus: "It has not seemed to me inconceivable that one should have an obligation and recognize that one has it and yet have no motivation to perform the required action" (Frankena 1958, pp. 42 – 43).

at all' or 'not caring very much' or 'caring less' about moral requirements (Zangwill 2008, p. 101). Zangwill rightly pointed out that we must not present indifferent agents as people who reject morality. The temptation of painting moral indifference as rejection of morality is seen in the case of *amoralism*. On the one hand, some externalist might think that it is only such a strong position of amoralism that guarantees a decisive counterexample to internalism. On the other hand, internalism seems to attack externalist cases of indifference from the perspective of rejection of morality, thinking that such a position flies in the face of categoricity of moral requirements. However, we are not claiming that moral demands do not apply to indifferent agents. Rather it is argued here that agents are not motivated by them because they do not care enough about moral issues (we shall return to this issue in the next chapter). That said, given that we do not need to construe indifference in such a strong term – complete indifference, we shall take it as "*the phenomenon of not caring very much about the demands of morality*" (Zangwill 2008, p.101). Following Zangwill's framework of indifference, our goal is to argue that people's interest, care or desires come in various strengths; and that indifference is actually possible because people care varyingly about moral issues. Assuming this argument works, then it serves as a counterexample to the internalist motivational efficacy claim of RMI**.

Zangwill's indifference argument is premised on the idea of degrees of beliefs as well as strengths of desires. He rightly pointed out one of the often overlooked driving elements of moral motivational debate is the "Proportional Determination Thesis", the view that the "strength of moral desire is proportionately determined by degree of moral belief" (Zangwill 2008, p. 95). The internalist claim implies, among other things, that motivation is essential to moral beliefs, that is, they motivate efficaciously. Given this, it is argued that if two persons are alike in their moral beliefs, it is necessary that they will be motivated alike given the claim that moral beliefs are motivationally efficacious independent of any additional desires. In other words, it is not a matter of contingency that motivation follows directly given that their moral beliefs are alike in every respect. It would only amount to inconsistency should the internalist claim that the content of belief of one of the persons is motivational efficacious, whereas the other not. If

motivation is essential to moral beliefs as internalist claims, then the content of moral beliefs of agents with equal cognitive dispositions *must* motivate them alike. On the contrary, it is actually possible for agents to share similar cognitive states, dispositions, beliefs and yet motivationally respond differently. Consider the Augustine's example in *De Civitate Dei*:

> Suppose that two men, of precisely similar disposition in mind and body, see the beauty of the same woman's body, and the sight stirs one of them to enjoy her unlawfully, while the other continues unmoved in his decision of chastity. What do we supposed to be the cause of an evil choice in the one and not in the other? What produced that evil will? …The mind? Why not the mind of both? For we assumed them to be alike in both mind and body […] What other reason could there be than his will, given that their dispositions were precisely the same, in body and mind?

An agent might hold a genuine moral belief, but if he does not care about the desirability of the belief that he ought to do the action, he will not be motivated by his moral belief (we shall return to this claim in detail in chapter 6). In other words, given the different intensities of individual's care about moral issues, it is possible that the phenomenon of indifference might occur between two persons sharing similar moral beliefs. More so, Zangwill argues that indifference can as well "be a matter of a person ceasing to care as much as he used to while his moral beliefs remain unchanged. Or it might be the possibility that a person at a time cares less than he actually does at that time while moral beliefs remain constant" (Zangwill 2008, p. 101).

4.6 Explaining Olivia's Behaviour

Olivia exhibits features of indifference: She does not seem to care, at least, about the moral issue at stake. Although, she is capable of forming and holding genuine moral judgements, she remains unmotivated or unmoved by them. We can attempt explaining her behaviour based on the two main categories outlined by Zangwill, namely the trans-personal and trans-temporal cases of indifference. In the former case, recall the incident

between Olivia and Emma. Both share the moral belief that environmental harm is morally unjust. However, while Emma was motivated, Olivia remained unmoved in the face of the same moral belief. The internalist thinks that her behavior is odd given that motivation is essential to moral beliefs as well as the fact that their moral beliefs are alike. However, Olivia's behaviour is not odd. It is actually possible that the strengths of her interests or care about moral issues vary. To illustrate this, imagine that Olivia was once highly active and took part in various environmental actions. However, recently she experienced that all their efforts made no (substantial) difference at all. Increasingly, her motivation to engage in such actions starts to dwindle, although she still strongly believes that the cause is morally right and even warrants actions. Now, she is completely worn out act accordingly.

Furthermore, assuming we rule out the cases of errors related to cognition and applications of moral concepts; and that they share precisely similar dispositions in mind and body. It is possible that she was not moved because not of her moral belief was less genuine than that of Emma, but because she does not care very much about the moral issue in question or moral demands in general. As we illustrated above, it is possible that she once cared about such actions, but now such a motivation is longer there. Given this, it might be claimed, contrary to the internalist claim, that:

(1) If agents A and B judge that φ is morally required, it is possible for A and B to be motivated differently (hence, not necessarily to φ) given their respective degrees of care about φ, while their moral belief φ remains unchanged.

In the latter case, namely, the trans-temporal case of indifference. Suppose Olivia used to care about morality, but of lately she started caring less about moral issues. (It might as well be that she cares about moral issues, but of lately she started caring not very much about environmental matters related to morality). On this level, her care about moral demands has become less than usual. As in the first case, she not only grasps the content of moral belief, but also she genuinely believes that environmental harm is morally unjust and yet she has no motivation to sign the petition. Given this, it might be claimed, contrary to the internalist claim, that:

(2) If an agent A judges that φ is morally required, it is possible for A not to be motivated given a change in her care about φ, while his moral belief φ remains unchanged.

Notice that in both cases that Olivia did not completely reject moral demands, at least, she continues to hold her moral beliefs. Notice also that other concerns did not matter more to her than morality. In other words, she is indifferent to her moral belief, because her care about the moral issue in question is not proportionately determined by the degree of her moral belief. Alternatively, it is possible that a change in Olivia's moral belief will not necessarily provide a change in her care about a new belief as the internalist claims. This is because we seem to care more or less about morality regardless of the genuine contents of moral beliefs we hold. In sum, Olivia might share moral beliefs with the rest of us, but if she cares less, she will be indifferent to morality. So also, she might have cared about morality (like the rest of us), but if she cares less now than usual, she will be indifferent to the moral demands that she used to care about, while her moral beliefs remain unchanged.

Nevertheless, it might be argued that moral beliefs and caring to act accordingly do not come apart. That is, to believe that X is morally required is inevitably to be motivated to X. Given that moral beliefs are taken to be best practical judgements of reason; it is argued that agents cannot fail to be motivated by what they judged as morally required. However, do we necessarily adhere to (even the best of) our moral judgements? The mere fact that we want certain things more than others or believe certain thing to have greater degrees than others, if only roughly, seems to show, at least, the possibility of caring less about what we judge as good. In other words, denying this possibility seems to amount to the following claim, namely, 'to believe something is necessarily to care about it'.[63] Such a denial is problematic, for it might place the idea of moral agency under a grave risk. If agents lack the possibility of choosing freely, morality would become a suspicious enterprise. In addition, such a move might lead to determinism, the sort that eliminates the possibility of freedom to choose. It is against this background that Henry

---

[63] On the contrary, belief is different from caring, the former is cognitively-laden, whereas the latter is an emotional capacity as I will argue in chapter 6.

of Ghent in his *Quodlibet* argues thus: "We must assume that [there is] over and above the freedom in reason to judge [*libertas arbitrandi*] and there is in the will a freedom to choose what is judged [*libertas eligendi arbitratum*], so that the will does not choose with any necessity even what reason judge after deliberation"[64] (Henry, *Quodl.* 1. 16, 5:102; Hoffmann, 2008).

Olivia has the possibility of (freely) choosing to care more or less about moral issues. She can as well choose not to care as she used to in the past. The case of indifference, hence Olivia's case, is actually possible given that people have the possibility of choosing to care or not to care at all; and there are cases where people freely decide to be indifferent to moral issues. Consider one of the three examples presented by Zangwill:

> [A] mercenary I once met on vacation exuded moral indifference. He was in control, reflective and articulate. Everything he said convinced me that he was perfectly aware that his vocation was genuinely morally wrong, not merely what people conventionally call 'wrong'. He fully understood the wrongness of his vocation. Nevertheless, he was not very concerned about that. He was more concerned with his immediate interests and concerns, that is, colloquially, looking after number one. There was no moral cognitive lack. He made that quite clear. Indeed, he insisted on it. The mercenary was unusually indifferent to the demands of morality; but he shared moral beliefs with the rest of us, and with his former self. He insisted on that (Zangwill 2008, p. 102).

Like Olivia, the mercenary in Zangwill's example is not suffering from psychological impairments. He knows fully well that his 'vocation' is morally wrong, hence knows what morality demands, but he freely chooses to be indifferent to those demands. In fact, he freely chose this vocation. Cases of moral indifference are part of our ordinary moral

---

[64] Super libertatem ergo arbitrandi in ratione oportet ponere libertatem eligendi arbitratum in voluntate, ut voluntas nulla necessitate eligat etiam quod ratio sententiat. . " (*Quodl.* 1. 16, 5:102; Hoffmann, Tobias (2008). Henry of Ghent's Voluntarist Account of Weakness of Will. In *Weakness of Will from Plato to the Present*. Catholic University of America Press).

experience; and Foot elegantly captures the possibility of this phenomenon in the following words: . . . one [*can*] be indifferent to morality…. (Foot 1978, p. xiv)

The emphasis is on *can* – normal people *can* freely choose to reject morality or care more or less about moral issues. We can be indifferent.[65] In the opening chapter, Olivia asks her interlocutor:

Do you mean, I've reasons to?

I see it not...

She concludes:

Even if I do,

I care not about it.

4.7 Conclusion

The case of Olivia seems to show that it is possible to hold moral beliefs and yet not be motivated accordingly. In other words, the claim that moral judgement is motivationally efficacious is false. Thus, since robust internalism seems not to present a plausible solution, in the next chapter we shall consider whether or not conditionalized version of realist internalism offers a better solution to the moral motivational problem.

---

[65] The fact that we can decide against the best practical judgement of reason explain why we can actually desire the bad. In his work '*Desiring the Bad: An Essay in Moral Psychology*' Michael Stocker argues that this reality is part of our normal moral experience. Stocker argues that cases where people fail to do or act according to their best decision or intention are in fact part of our common moral experience (Stocker 1979). In essence, he argues that it is not necessary that moral judgement of what one believes to be good or best or right necessarily motivates. Stocker writes that: "Motivation and evaluation do not stand in a simple and direct relation to each other, as so often supposed. Rather, they are interrelated in various and complex ways, and their interrelations are mediated by large arrays of complex psychic structures, such as mood, energy, and interest" (Stocker 1979, 738 – 9).

Chapter 5

'The Moral Problem': Smith's Solutions?

> There is not a thing, rationality,
>
> with a nature waiting to be investigated.
>
> Rationality is a norm-relative affair.
>
> **Charles Pigden**, *Hume, Motivation and the Moral Problem,* **2007**

5.0 Introduction

Unlike the robust realist internalism, it is argued that adopting a moderate position would not defeat the internalist project. On the contrary, it enables the internalists to paint a more plausible picture of moral motivation. On this level, the task that lies before this group of internalists is that of specifying those conditions under which the *necessary* connection between moral belief and motivation holds. Nevertheless, if this version of internalism is to be taken seriously it has to capture the features of internalism without trivializing either the essentiality claim or the necessity claim of internalism. Given this, the emphasis shifts from robust internalism [RMI (as we saw in the last chapter)] to *Conditionalized Realist Internalism* (CRI) the latter of which holds thus:

> (CRI): Necessarily, if agents judge or believe that they are morally required (or morally ought) to φ; they are to some extent motivated to φ in case they are in the condition 'C'.

In what way(s) then are we to understand the 'C' in CRI?[66] In their work *Recent Work on Motivational Internalism,* Björklund et al. (2012) outline three main ways of characterizing the conditions. First, *Psychological Normalcy*: It is argued that internalism

---

[66] Caj Strandberg observes the following:

> I take this restriction to mean that C cannot be understood in such a way that it becomes trivially true that a person who satisfies C is motivated in accordance with her moral judgement. Thus, C cannot be understood to simply state that a person who satisfies it is motivated in accordance with her moral judgement. Similarly, it cannot be understood as a mere *ad hoc* condition, such as a negation of a disjunction of mental conditions which is thought to result in the fact that a person is not accordingly motivated. (Strandberg 2013, p. 26)

is plausible insofar as the agents are psychologically normal. The normalcy meant here is the absence of psychological conditions such as depression, apathy, weakness of the will, spiritual exhaustions, etc. To illustrate this view,

> PN*CRI: Necessarily, if agents judge or believe that they are morally required (or morally ought) to φ; they are motivated to φ in case they are *psychologically normal.*

Second, *Practical Rationality:* This view ties internalism to the requirements of practical rationality. Here, being rational means being motivated to do what one morally believes; and as such, motivational failure is considered as a case of irrationality. The psychological conditions listed above are considered here as cases of irrationalities (at least, Smith thinks so). To illustrate,

> PR~CRI: Necessarily, if agents judge or believe that they are morally required (or morally ought) to φ; they are motivated to φ in case they are *practically rational.*

Finally, *Moral Perceptivity:* This version of conditionalized internalism is related to the view we saw in the previous chapter. That is, if any agent fully grasps or cognizes moral properties of actions or moral obligations, she is motivated. Since some of the objections raised against platonic internalism (namely, this form of internalism requires agents to possess some sort of special moral perceptive faculties or mysterious moral endowments) apply here as well, I will drop this position. Hence, for the purposes of this chapter, I will be dealing with PR~CRI. This position seems very attractive because it claims that (1) rationalism entails internalism; and (2) realist internalism is compatible with desire and means-end belief, thus espousing Humean psychological conception of motivation. (3) In addition, the condition of PR~CRI is linked to the condition of PN*CRI, namely psychological normalcy. We assume that a (practically) rational agent is at least psychologically normal. As we shall see later, one of Smith's requirements for a reliable connection between moral judgement and motivation is the absence of psychological impairment.

The idea driving PR~CRI is the notion of conceptual truth (Korsgaard 1996; Nagel 1970; Smith 1994). Since Smith's position tersely captures both rationalism and internalism, the goal of this chapter is to ascertain whether his argument for conditionalized internalism gives us reason to abandon cognitivist externalism. Smith has two internalist solutions namely, the argument from Rationalism and from Practicality Requirement respectively:

(a) *Rationalism:* If is right for agents to φ in circumstances C, then there is a reason for them to do φ in C.

(b) *Practicality Requirement*: If an agent judges that it is right for her to φ in circumstances C, then she is motivated to φ in C provided she is practically rational.

Generally, Smith's internalist position, as I understand it here is based on his confidence on the conceptual truth of rationalism (Smith 1994, p.65). Roughly, according to him, since it is a conceptual truth that if an agent judges that she has a reason to do something, it would only amount to irrationality if she fails to perform it. Alternatively, if an agent judges that it is right to φ and he is rational, then it follows that he is motivated to φ *ceteris paribus*. Notice that if an agent is rational that is, if he has a reason to φ and he fails to be motivated to φ, then he must be practically irrational. The claim seems to show that it is possible to be rational without being practically rational. Given this, we shall proceed by considering Smith's argument from Rationalism. First, our strategy is to refute his rationalist claim by showing that his account of normative reason is based on an inconclusive model of the norm of rationality. Second, I argue that his argument from practicality requirement, thus his attack against externalism, fails. In other words, we have reasons to adopt the externalist theory of motivation.

5.1 Moral Rationalism

Thomas Nagel's claim is the *locus classicus (*at least in the recent debate) of the version of moral rationalism, the type defended by Smith and the like. In *The Possibility of Altruism*, he argues thus,

> Just as there are rational requirements on thought, there are rational requirements on action…If the requirements of ethics are rational requirements; it follows that the motive for submitting to them must be one, which it would be contrary to reason to ignore. (Nagel 1970, p. 3)

Nagel argues that both thought and action adhere to the constraints of reason. Thought (here a moral one), on this construal, is in the business of expressing beliefs. Rationality, among other things, requires that our deliberations and ultimately our beliefs, capture the objective (moral) state of affairs. Precisely, moral thought has a rational embeddedness. So also, the requirements of moral actions are requirements of reason. However, how does this come to be? Concretely put, where does this intuition come from? Smith provides us with the following answer:

> Our concept of a moral requirement is the concept of a reason for action; a requirement of rationality or reason. (Smith 1994, p. 64)

Notice that rational requirement is based on some sort of conceptual claim, namely, the kind that ties requirement to the meaning or truth of concepts. Smith claims the following: by understanding the concept of moral requirement (that is, what he is morally required to do), an agent grasps the requirement of reason for action as well. In other words, moral demands are intrinsically embedded in the requirements of rationality or reason. Let us, following Richard Joyce (2002), and Shaun Nichols (2002), call this view *Conceptual Rationalism* in order to distinguish it from related views.[67] Roughly, to illustrate,

---

[67] The other views are psychological and justificatory rationalism. The former conceives moral judgement as a sole product of rational faculty. On this level, it is argued that our moral judgement causally flows from the rational faculty (Kohlberg 1976; Colby and Kohlberg 1987; Piaget 1933). The latter grounds the justification of morality on reason. The justificatory rationalist believes that moral explanation is a rational explanation.

Conceptual Rationalism (CR): If it is right for an agent to do something in circumstances C, then there is a *reason* for him to do it in C.[68]

If we take *reason* in CR simply to mean what counts *for* or *against* doing something, then this construal will accommodate varieties of reasons that will undermine the rationalist understanding of both moral authority and internalism. In other words, the moral rationalist will not easily concede to the possibility of other 'foreign body' *reasons* figuring in moral demands. This is because such a concession might open up the possibility of explaining moral motivation in terms of other 'reasons' external to moral reason. To block this objection, let us follow the rationalists in calling reason in CR a *normative reason*.[69] Hence, assuming that the normative reason in question has appropriate moral content, we can illustrate conceptual rationalism as follows:

(CR*): If it is right for an agent to do something morally φ, then there is a normative reason for him to φ.

5.2 Smith's Notion of Normative Reason

The notion of normative reason is crucial to Smith's internalist solutions. To this end, his goal is that of offering a theory that explains "about those of our normative reasons which express norms of rationality or reason" (Smith 1994, p.131); and how it (that is, normative reason) provide agents grounds for moral action. Smith is aware that such a task is riddled with the ambiguity of concept. Precisely, reason could be confusing, because it is used both as *normative (reasons for action)* and *motivating* reason. Smith argues that it is false to claim that *all* reasons are explainable both in "an explanatory and justificatory

---

[68] Based on Mackie's distinction of conceptual and substantive questions (the former asks, 'What is our concept of x?' Whereas the latter asks, 'Is there anything like x?'), Smith wants us to understand his form of rationalism as a conceptual claim and not substantive claim. On this reading, he is concerned with what our concept of moral requirements is: "Taken in the first way, rationalism is a claim about the best analysis of moral terms. Taken in the second way, rationalism is a claim about the deliverances of the theory of rational action" (Smith 1994, p. 65).

[69] Even referring to it as normative reason does not entirely solve the problem because normative reason can be prudential, hence non-moral in nature. For this sake, the normativity of moral reason must have the right (substantial) content that differentiates it from other non-moral normative reasons.

dimension"[70] (Smith 1994, p. 95). However, this is not to be understood as a denial of those cases where motivating and normative reasons seem to intercept each other in the agent's action (as we briefly noted in the previous chapter). On this point Smith makes the following observation:

> Motivating and normative reason do have something in common in virtue of which they both count as reason. For citing, either would allow us to render an agent's action intelligible. This is essential. For there is an *a priori* connection between citing an agent's reasons for acting in a certain way and making her acting in that way intelligible: that is, specifying what there is to be said for acting in the way in question. (Smith 1994, p. 95)

Nevertheless, according to Smith, this is not to suppose that there is no difference between them. In fact, his claim is that once we establish a clear distinction between them, we shall come to grasp why anti-Humean psychology is a failure.[71] Alternatively, such a distinction, to Smith's mind gives us a better strategy for attacking Hume. That is, it specifically makes us see what was actually wrong in Hume's motivation theory, namely, his construal of normative reason, rather than his motivating reason. How, then, does Smith understand the difference between motivating and normative reasons?

According to him, we can explain the claim that 'A has a reason to φ' in two ways depending on where our emphasis lies. It can either be normative or motivating reasons to φ. In the case of the former, to say that 'A has a reason' amounts to saying that "there is some normative requirement that she φ's, and is thus to say that her φ-ing is justified from the perspective of the normative system that generates that requirement" (Smith 1994, p.95). Notice that A's φ-ing figures in the justificatory network. In the case of the latter, when we say that 'A has a reason to φ' we mean something different from the

---

[70] Smith accused Woods (1972) of holding this position: "Michael Woods argues some time ago that 'the concept of a reason for an action stands at the point of intersection, so to speak, between the theory of the explanation of actions and the theory of their justification'. Woods's idea was that reasons – all reasons – have both an explanatory and a justificatory dimension. I doubt very much that that claim is true…." (Smith 1994, p. 95)

[71] Nagel, McDowell and Platts were his main targets. He thinks that they wrongly rejected Hume's belief and desire construal of moral motivation because they failed to establish the difference between Hume's account of normative and motivating reasons.

justificatory feature of A's φ-ing. Here it is the state of A, precisely her psychological state that is at stake. According to Smith, this state *explains* A's φ-ing; and "given that an agent who has a motivating reason to φ is in a state that is in this way potentially explanatory of her φ-ing, it is thus natural to suppose that her motivating reason is itself psychologically real" (Smith 1994, p. 96). Unlike the justificatory case of normative reason, the motivating reason enters into an explanatory network of A's φ-ing. In other words, Smith wants us to understand normative reason as *truths* or *propositions* about what A 'should' or 'ought to' do; and motivating reason as the 'psychological states' that explains the φ-ing of A. The former justifies, whereas the latter explains.

Furthermore, Smith thinks that Hume's mistake was that of integrating a faulty account of normative system into his psychology of human motivation. He is confident that once the error is uprooted and replaced with a plausible account of normative reason, then Hume's desire and means-end belief would become a powerful theory for explaining human motivation,[72] hence a tool for solving the 'moral problem'. Therefore, it was not Hume's construal of motivating reason that was wrong, but his account of normative reason. However, what was wrong with his normative reason? For Hume, the claim that 'A has a reason to φ' means that φ is whatever counts most for A's desire. Smith argues that Hume's account of normative reason makes 'reason' responsive to contents of agents' desires. In other words, normative reason is relative to whatever each agent thinks.

Contrary to Hume's account, Smith opens his account of normative reason by considering possible perspectives of interpreting agent's *intentional action*. According to him, we can approach such actions from two perspectives. On the one hand, we can explain intentional action by looking at its end, outcome or cause. Notice that this perspective seems to fit into Hume's model of the belief-desire pair, hence, the motivating reason described above. On the intentional interpretation, therefore, we can explain an agent's action by locating her desire and matching it with her respective beliefs. On the other hand, we can explain intentional action from the deliberative perspective. As the

---

[72] Smith's goal is to show that "Humean theory is the expression of a simple but important truth about the nature of motivating reasons, a truth that anti-Humeans have failed to appreciate either because they have failed to distinguish motivating reasons from other sorts of reasons, or because they have an inadequate conception of desire, or because they have overlooked the implications of the fact that reason explanations are teleological" (Smith 1994, p.93)

name suggests, intentional action is approached from its deliberative patterns. Here, what counts as rational is that which the agent considers as desirable. Precisely, in the process of deliberation, there are grounds that count more than others. Thus, it is argued that they are supposed to provide agents with reasons for acting in certain ways. Smith argues that such grounds constitute rational justification for action – normative reason (Smith 1994, p. 131). Notice that these perspectives are alternative ways of describing the distinction between motivating and normative reasons. However, what does Smith's normative reason amount to then?

He argues that normative reason lies on agents' value attitudes. The notion of value plays a crucial role in his account of normative reason. For him, it is not enough to *desire* that we have a reason to act in a certain way; rather we have to *value* it. That is, specifically, we have to value acting in accordance with our reason. Further, Smith dismisses desire as a feature of agents' value-content on the grounds that agents can "desire to act in a certain way without valuing acting in that way; without thinking that so acting is rationally justified" (Smith 1994, p. 134). To his mind, value is a belief. But if this is true, then, how does belief motivate an agent in the absence of desire in the agent's value-content? Once again let us recall Hume's constraint on beliefs – belief alone cannot motivate. If Hume is right, how then does Smith's account of normative reason, which is essentially belief, constitute motivation?

Now Smith falls back to some sort of *a priori* claim, namely, a conceptual connection between belief and desire. He argues that agents rationally desire what they value, all thing being equal. For him, it is conceptually true that,

> If an agent believes that she has a normative reason to φ, then she rationally should desire to φ. (Smith 1994, p.148)

However, notice that this conceptual truth does not tell us what is meant by 'having a normative reason'. Regarding this concern, he argues that an agent has a reason for acting in a certain desirable way if she were to be rational, that is, were she to conform to the 'norms of rationality'. Hence, following Christine Korsgaard, he strongly believes that

119

this conceptual truth explains the connection between belief that something is a reason for action (that is, the desirable) and desiring it,

> For it is a platitude to say that what it is desirable that we do is what we would desire to do if we were fully rational; that what we have normative reason to do is what we would desire that we do if we were fully rational. (Smith 1994, p. 150)

5.3 Smith's Argument from Rationalism

Based on this account, believing that something is desirable amounts to desiring it; and as such, agents have reasons, namely normative reasons, to do what they desire *if they were fully rational.* When we judge that it is desirable, we are expressing our belief about our reason for action. Hence, "our moral judgements are an expression of our beliefs about what we have normative reason to do" (Smith 1994, p. 185). Suppose an agent, A, judges that it is right to φ, Smith claims that, given that A's judgement (that is, her believing that she has most moral reasons to φ), she is motivated to φ if she were fully rational. This thought is grounded in the following claim:

> It is a platitude that an agent has a reason to act in a certain way just in case she would be motivated to act in that way if she were rational. And it is a consequence of this platitude that an agent who judges herself to have a reason to act in a certain way – who judges that she would be so motivated if she were rational – is practically irrational if she is not motivated to act accordingly. (Smith 1994, p. 62)

To illustrate,

> CR**: If any agent judges that it is right for her to do φ, she is motivated to do φ if she were fully rational.

The 'rightness' in CR* is not to be understood as what motivates; rather it is the judgement that 'it is right', hence, it is in believing that one has normative reason to do φ in C. Smith believes that the constraint of CR** applies to *all* agents, because

(1) "Moral requirements apply to rational agents as such" (Smith 1994, p.85).

(2) However, why is it so? This is because "we think of the moral requirements that apply to agents as themselves categorical requirements of rationality or reason" (Smith 1994, p.85).

Claims (1) and (2) are motivated partly by Kant's influence on Smith, and partly by Hume's account of normative reason. Smith believes that given that moral requirements are rational requirements, they apply to all rational agents. They bind any agent that has a reason to do φ in C; hence, adhering to the requirements of morality is a categorical imperative of rationality. In this case, normative reason ceases to be responsive to agents' desires. Assuming Smith's argument sticks, then Hume's account of normative reason is false: "The non-relative conception of normative reasons, this claim has a straightforward truth condition: it is desirable that p in C just in case we would all desire that p in C if we were fully rational" (Smith 1994, p. 166). Further, he argues

(3) If (2), then "it is a conceptual truth that if agents are morally required to act in a certain way then we expect them to act in that way" (Smith 1994, p.85).

(4) Then, "being rational, as such, must therefore suffice to ground our expectation that rational agents will do what they are morally required to do. For the only thing we can legitimately expect of rational agents as such is that they do what they are rationally required to do" (Smith 1994, p.85).

These assumptions form Smith's core argument for rationalism. However, in what follows, I argue that this argument is fraught with a number of difficulties. It is not clear whether (1) is plausible, because assuming there are cases where being rational does not entail being moral, then claiming that moral requirements are rational requirements and apply to agents as such would be false. The *amoralist* seems to present us with a case where being rational and being moral come apart, at least on the conceptual level. It is possible to imagine an agent who judges that 'it is right to φ in C', but still can't see reason why she has to place herself under the requirement of morality to φ in C. David Brink strongly believes that amoralism is actually possible (Brink 1989, pp. 148, 45 - 50). If such cases were not possible, it would be senseless asking: *why be moral?* in the first

place. If it were always the case that moral requirements apply to rational agents as such, then it would be 'one thought too many' to justify why rational agents have to be moral or why we expect rational agents to act in accordance with moral requirements. However, there are cases where agents exhibit full rationality (even fulfilling Smith's constraints of being rational) without being moral or accepting the requirements of morality. Further, suppose our amoralist friend upon proper deliberation comes up with the belief that it is wrong to x in C (that is, his belief captured all the features required to make the rationalist moral judgement right), she can still ask: why do I have to be moral? Olivia asked the same question at the beginning of chapter 1. In other words, '*Why (should I) be moral?'* seems to point to the fact that being rational does not entail being moral; and if Smith denies this fact, then he would be begging a question against the cases of our amoralist friend, Olivia and the like.

Smith's assumption would work assuming that the rational agents are all moralists: that is, agents that believe that morality is worth caring for. Even on this construal, moral requirements would only apply to them loosely. This implies that moral requirements would apply to moral egoists[73] as well as altruistic moral agents, deontologists, as well as utilitarians. However, Smith would be begging the question against other rationality models of morality, should he claim that there is only one model of rationality. It is possible that positions such as moral egoism or utilitarianism can fulfill the constraints of being fully rational. After all, Nagel holds that altruism is one of the rational requirements on moral action. Hence, there are other possible rational requirements; and our moral experiences show us that people (i.e. fully rational agents) actually act on the basis of self-interest, maximization of desires, altruism, etc. It is open whether it is irrational to prefer one's interest to other people's interests. Hume argues that:

> 'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. 'Tis not contrary to reason for me to chuse my total ruin,

---

[73] Here it is moral egoism (the view that agents' moral motive for acting is based on their own self-interest) that is meant. This position is to some extent related to rational egoism (the view that agents' motive for acting is rational if and only if it maximizes agents' self-interests). This position is not to be confused with psychological egoism (the view that agents are motivated only to act in their own self-interest).

to prevent the least uneasiness of an Indian or person wholly unknown to me. 'Tis as little contrary to reason to prefer even my own acknowledg'd lesser good to my greater, and have a more ardent affection for the former than for the latter. (THN 416)

If Hume is right, that is, if it is not irrational for an agent to prefer her own interest to those of others (if it is not irrational for an agent to not prefer what the other person considers as most important reason for action (or greater good) over her own acknowledged 'lesser' reason for action), then it is not irrational for her to build her moral requirements on other features. In other words, assuming that moral requirements are rational requirements, that is, applying to rational agents as such, the moral egoist and utilitarian would be moral because they are rational.

Furthermore, how true is the claim in (2)? Suppose Smith is right, that is, that "we think moral requirements that apply to agents as themselves categorical requirements of reason" (Smith 1994, p.85), then such requirements are supposed to apply to *all* agents having normative reason when confronted by the same circumstance C, *if they were fully rational*. On the contrary, I doubt the plausibility of this claim. Assuming that moral requirements are categorical imperatives, Smith's account of normative reason fails to meet the conditions of such imperativity. Generally, imperatives are propositions, which command some sort of "should" or "ought" to be done. Hence, imperatives have the "force of a command of reason" according to Kant. According to Kant, categorical imperatives are relevant to morality, because unlike hypothetical imperatives, they universally and unconditionally apply to rational agents irrespective of circumstances, inclinations, desires and the like. Hence, the thought that 'it is right to do φ', on the construal of categorical imperative does not depend on *if* the agents have desires to φ. However, the purported moral requirement of Smith's categoricity is *conditional*. That is, it is founded on a counterfactuality, namely on a platitude that agents have normative reason to act *if they were to fully rational*. According to Smith:

[P]latitude tells us that what it is desirable for us to do is what we would desire that we do if we were fully rational. In other words – and now we are turning the

platitude into an analysis, for we are making explicit distinctions that are at best only implicit in the platitude – it tells us that what is desirable to do in certain circumstances – let's call this the 'evaluated possible world" – is what we, not as we actually are, but as we would be in a possible world in which we are fully rational – let's call this the 'evaluating possible world' – would want ourselves to do in those circumstances. That is, it tells us that facts about the desirability of acting in certain ways in the evaluated world are constituted by facts about the desires we have about the evaluated world in the evaluating world. (Smith 1994, p. 151)

Notice that Smith's categoricity of rationality applies only if the agents were to consider themselves to be fully rational. However, it is not clear how 'if we were fully rational' is to be understood. That is, is the *if-clause* (an agent has reason to desire φ *if she were rational*) supposed to satisfy the unconditional feature of categorical imperatives? It seems that the *if-clause* conditionalizes the applicability of categorical requirements of rationality. In other words, the moral demand of categorical requirements of rationality applies to agents only if they were fully rational. I read Kant's categoricity to mean that morality is inescapable because the force of categorical imperatives applies to agents unconditionally. However, it is unclear how Smith's categoricity, which applies to agents on the if-clause basis is supposed to capture the unconditionality and universality features.

More so, Smith's conception of categoricity leaves open what it means for agents, even when they converge, to be fully rational. It is possible that a group of agents agrees that 'it is right to φ*' and thus, 'desire to do φ*' assuming they are fully rational. Suppose another group comes up with the position that 'it is right to φ' and thus, 'desire to do φ' assuming they are also fully rational. Both groups have legitimate normative reasons and full rationality respectively, all things being equal. Thus, it is difficult to churn up justifications why we should prefer φ* to φ. To flesh out this claim, let us assume that the first group espouses *hard paternalism*, whereas the second group holds *soft paternalism*. Suppose that upon fulfilling Smith's requirements of full rationality, the former converges that it is right for physicians to fully decide for the patients' values and interests because they are competent and know what is best for them. Whereas the latter thinks the

124

physicians, can only step in and help the patients to realize their values and interests without eliminating their freedom of choice. All things being equal, upon fulfilling the requirements of being truly rational, it could be argued that both groups have normative reasons to φ* and φ respectively. In other words, by claiming that all agents would have a normative reason to do something if they are all in the same circumstance begs the question against agents' varying degrees of valuing things. Precisely, Smith's claim apparently presupposes a reductionist and monistic account of value, which implies, against the notion of value pluralism, that there exists always one objectively valid value judgment. However, it is a part of our moral experience that people value things differently. Value pluralism maintains that "fundamentally human values are irreducibly plural and 'incommensurable', and that they may, and often do, come into conflict with one another, leaving us with hard choices" (Crowder, 2002, p. 73. For discussions see Berlin 1956, 1990, 2000; Gray 1995b, 1998, 2000; Raz 1986; Stocker 1990; Kekes 1993; Richardson 1997; Chang 1999; Galston 2002).[74]

Similarly, Richard Joyce (2001) observes that two fully rational persons (Ernie and Bert) might have normative reason to do φ, but differ in terms of desiring φ (Joyce 2001, p. 78). Notice that if rationalism entails internalism as Smith claims, both Ernie and Bert would be motivated to φ since they have normative reason to φ. However, it is actually possible that they see reason in φ, but their varying degree of valuing φ might not give one of them equal reason for acting on φ. Further, suppose for the sake of argument that Ernie and Bert have normative reason expressed in the belief that 'it is right to φ', suppose also that they have similar content of desire, it is also possible that the degree of their valuing φ might influence their reason to do φ even in the case of full rationality. Thus, if it is the case that fully rational agents actually differ in their desires as well as degree of desiring, then it is not true that 'If all agents judge that it is right to

---

[74] One should not confuse value pluralism with some sorts of subjectivism or relativism. Although values are multiple, they are, however, objective, because they *always* refer to something pertaining to value notwithstanding its various manifestations in, for example, cultures, societies. Isaiah Berlin argues that "the multiple values are objective, part of the essence of humanity, rather than arbitrary creations of men's subjective fancies" (Berlin, 2000a, p. 12; for discussion on the difference between pluralism and relativism see Berlin, *Alleged Relativism in Eighteenth-Century European Thought*, in Berlin 1990; Crowder 2004, pp. 114 -123).

do φ in C, then there is a reason for them to do φ in C'. In addition, to making the idea of categoricity of moral reasons responsive, that is, relative to whatever the agents' desires to do upon fulfilling the requirements of rationality, we shall see below why Smith's account failed to capture the unconditional features of categorical imperatives.

5.4 Philippa Foot's Charge against Categorical Imperatives

In *Morality as a System of Hypothetical Imperatives* (1972), Philippa Foot offers a more extended charge against the categoricity claim of moral requirements as rational requirements. Recalling the distinction between hypothetical and categorical imperatives above, Foot asks, "Is Kant right to say that moral judgements are categorical, not hypothetical, imperatives?" She answers thus:

> It may seem that he is, for we find in our language two different uses of words such as "should" and "ought," apparently corresponding to Kant's hypothetical and categorical imperatives, and we find moral judgements on the "categorical" side. (Foot 1972, p. 307)

In her analysis of the uses of "should" in hypothetical and categorical statements, she notes that Kant's use of "should" in moral statements goes beyond mere linguistic application. Kant embeds the "moral-should" on some sort of metaphysical dignity and necessity (Foot 1972, p. 308), such that, "You should not break your promise" binds us unconditionally. However, she observes that there are non-hypothetical and non-moral contexts where categoricity features apply. According to her, the non-hypothetical "should" applies to etiquette or clubs' rules as well. For example, teachers tell their pupils, "You should not disturb the class", regardless of their desires. If categoricity applies in the case of "You should keep your promise" regardless your interest and desire, she then asks, why should not categoricity apply to "You should not disturb the class"? Foot argues that there is no difference in the categoricity feature of both non-hypothetical statements, notwithstanding that one is moral and the other not. Foot's worry is thus, why do Kantians claim that moral categorical imperatives are reason-giving for action? Why not assign

reason-giving character to other non-hypothetical statements? Why insist that moral categorical imperatives enjoy special status?

> The difficulty is, of course, to defend this proposition which is more often repeated than explained. Unless it is said, implausibly, that all "should" or "ought" statements give reasons for acting, which leaves the old problem of assigning a special categorical status to moral judgement we must be told what it is that makes the moral "should" relevant different from the "shoulds" appearing in normative statements of other kinds. To say that moral considerations are called reasons is blatantly to ignore the problem. (Foot 1972, p.309)

The rationalist has the option of choosing between accepting that all "shoulds" are categorical, or explaining away the special normativity of moral "should". If he takes the first horn the old problem will disappear but that would imply that moral statements are no more special than other non-moral statements. They do not possess any reason-giving status more than any other non-hypothetical domain. Further, since no one would regard a pupil who breaks the class rule as being irrational, it follows that – if the rationalist accepts that "should" applies across other domains – then disobeying moral rules is not irrational. If, however, the rationalist takes the second horn, he has to provide justifications for the special status of morality. Foot considers the arguments from *coercion* and *feelings* as possible justifications for the special normative character of morality. However, she finds them wanting. According to her, we have no reason for conforming to moral judgements, because moral requirements are not rational requirements. In addition, assuming that not adhering to the requirements of morality is not irrational, then the amoralist and Olivia's behaviours are not irrational:

> The fact is that the man who rejects morality because he sees no reason to obey its rules can be convicted of villainy but not inconsistency. Nor will his action necessarily be irrational. Irrational actions are those in which a man in some way defeats his own purposes, doing what is calculated to be disadvantageous or to frustrate his ends. (Foot 1972, p. 310)

Foot concludes that moral requirements are not requirements of reason or rationality; and being indifferent is not irrational. Therefore, requirements of morality are not categorical imperatives. On the contrary, she advises us to regard them as hypothetical imperatives. In other words, Smith's assumptions in (1) and (2) are not true. That is, it is not true that "moral requirements apply to rational agents as such" and it is false to claim that "we think of the moral requirements that apply to agents as themselves categorical requirements of rationality or reason".

Let us consider whether assumptions (3) and (4) are plausible. Notice that they depend on (1) and (2). But given that we have found them to be false, I doubt very much whether (3) and (4) will succeed. Nevertheless, let us give them the benefit of the doubt. First, do we really expect agents to do what is morally required? Assumption (3) ("if agents are morally required to act in a certain way then we expect them to act in that way") is problematic because it is possible for agents to acknowledge that there is a moral reason to x, but lack the motive to act in the way we expect them. To this end, Smith modified (3) thus:

> We certainly expect rational agents to do what they judge themselves to be morally required to do: that is, we certainly believe not just that they should, but they will, other things being equal. (Smith 1994, p. 84)

The force of expectation is now tied to agents' judgement: 'it is right to do φ'. However, this argument works insofar as (1) and (2) are true: That is, we expect agents to do what they themselves judge to be morally required only if moral requirements are categorical imperatives. Moreover, since, as we saw above, that moral imperatives seem to be escapable (at least, in terms of not being motivated to act accordingly), it follows that our expectation on agents' moral behaviours is not unconditionally binding.[75] In other words,

---

[75]Miller elegantly argues that "[A]ll Smith is entitled to say is that we assume rational agents to judge truly on rational matters: to say that we assume that rational agents judge truly on moral matters is simply to assume exactly what the anti-Rationalist denies, namely, that moral matters are matters of reason. Again, Smith's argument clearly begs the question." (Miller 2003, p. 230)

Smith's confidence that agents will do what they are morally required to do lacks a solid justification. To this objection Smith replies as follows:

> (4) Being rational, as such, must therefore suffice to ground our expectation that rational agents will do what they are morally required to do. For the only thing we can legitimately expect of rational agents as such is that they do what they are rationally required to do. (Smith 1994, p. 85)

Given that agents are rational, we expect them to do what they are morally required to do, because they know and we also know that moral approval and disapproval are ubiquitous, Smith argues. Nevertheless, such purported moral approvals and disapprovals need to be supported by some sort of moral facts or at least moral conclusions realized by convergence of rational deliberations. Even at that, the matter is not fully settled, because ordinary experience shows us that "one man's meat is another man's poison" principle applies as well in the moral context. In addition, this fact seems to be ubiquitous in our moral lives: For example, what the pro-life members (they think abortion should be banned) disapprove might be what pro-choice member (they believe that individuals have autonomy with respect to their own reproductive systems, provided they do not breach other people's autonomy) approves. In other words, if being rational is the ground for our expectation; and assuming we cannot legitimately deprive either of the sides of their rationality claim, it is difficult to justify that claim 'we expect rational agents to do what they are morally required to do', where the expectation is based on our conceptions of rationality. That is, reasons that count *for* or *against* moral approval or disapproval respectively. In any case, (3) and (4) cannot succeed because (1) and (2) are false. In sum, Smith's argument from rationalism fails because moral requirements are not categorical requirements of rationality; and they do not apply to rational agents as such. Let us now turn to his second defense of internalism.

5.5 Argument from Practicality Requirement

In the argument from Practicality Requirement, Smith defends internalism, precisely, by attacking externalism. His argument opens with the following claim:

> It is a striking fact about moral motivation that *a change in motivation* follows reliably in the wake of a *change in moral judgement*, at least in the good and strong-willed person. (Smith 1994, p. 71)

The striking fact reflects some somewhat pairwise conceptual relations whereby a change in one order causes a corresponding change in another order. Smith argues that there cannot be a change in moral judgement without a corresponding change in motivation. Moral motivation supervenes on moral judgement and not vice versa. Further, Smith introduces the *ceteris paribus* clause: That is, the reliable connection presupposes [1] the absence of psychological impairments we listed above; [2] the absence of motivational features external to moral judgement; and above all, [3] agents must be good and strong-willed people. In other words, once these conditions are satisfied, Smith strongly argues that motivation necessarily and internally follows from moral judgement. To illustrate this point, suppose Pete believes that the government should have a full control of the internet and its contents. He believes that the government knows what is best for the citizens. However, after a long conversation with his girlfriend, he drops this position. Now he believes that even a gentle nudge from the government is an undue interference, hence a violation of the citizens' right to privacy. If Smith is right, in the wake of change of belief, Pete should be necessarily motivated to oppose government's control of the internet. Therefore, motivation reliably occurs in the presence of belief; and in the case of new belief, new motivation reliably occurs accordingly, thereby replacing the original belief and motivation (Smith 1994, p. 72). Smith then contends that internalism, as well as externalism, must explain this *striking fact about the reliable connection.* In fact, he is confident that only the internalist can explain it. However, notice that the reliable connection as it is presented above is a clear expression of internalism. Smith's explanatory proposal is problematic at least on two grounds. Let us begin by asking the

following: Is the reliable connection to be explained the one stated above? Alternatively, does he mean that both sides (namely internalism and externalism) have to (jointly) provide a reliable connection between moral judgement and motivation? If the former, then it is problematic: how does Smith expect the externalist to explain a reliable connection, when the connection in question is an internalist assumption? Notice that the reliable connection, as it was presented above, is not neutral, because it says that motivation follows directly from moral judgement; and this is exactly what the externalist rejects.[76] Just as it is weird to expect a moral non-factualist to explain a striking fact about the corresponding connection between moral beliefs and moral facts, so also it is strange for an internalist to expect an externalist to explain a reliable connection on his terms.

Further, since the striking fact about the reliable connection is an expression of internalism, it is counter-intuitive for Smith to invite the internalist to (re)explain what he has already established. That is, he already established that it is a striking fact, hence obvious truth, namely *a change in motivation* follows reliably in the wake of a *change in moral judgement* (at least in the good and strong-willed person). It is therefore unclear how the internalist is supposed to explain an established fact. If he means the latter, then Smith's accusation that externalism lacks the explanatory relations tools for establishing the reliable connection is unsound. In fact, the externalist has a more plausible account of moral motivation. Since I do not believe that there is such a thing as a reliable (where reliability means necessity) connection, I will turn down Smith's offer of explaining the reliable connection as it stands above. My objective in the next section is to argue that Smith's attack on externalism is unsustainable because his construal of internalism faces the same objections he raised against externalism.

---

[76]At least Smith does not mention (both in 1994 and 1996) that such moral judgement is supposed to motivate in connection with other factors. So it is disturbing ,for Smith, to claim that both internalism and externalism alike accept the premise, 'that it is a striking fact that when good and strong-willed people judge it right to φ in C they are motivated to φ in C' (Smith 1994, p. 175). It is an assumption gone wild, because it seems to compel the externalist to explain his account of reliable connection on an internalist premise. Hence, I think that his purported attack on externalism fails, because the premise is not compatible, in the first place, with externalism.

5.6 The Fetishist Argument

Smith's contention is that Pete's change in motivation follows internally from his new belief because he is a good and strong-willed person. Hence, it captures the internalist assumption that:

(1) If any good and strong-willed person judges that it is right to φ, then he is motivated to φ.

On this reading, an agent's motivation is solely tracked by just looking at his moral judgement. Smith advises us not to worry about any additional content of motivational states, because

> If an agent judges it right to φ in C, and she has not derived this judgement from some more fundamental judgement about what it is right to do in C, then, absent weakness of will and the like, the defender of the practicality requirement can insist that she will be motivated non-derivatively in φ in C. This is because, a non-derivative desire to φ in C in what her judgement that it is to φ in C causes in her. (Smith 1994, p. 73)

Smith introduces the non-derivative desire in the motivational activities of good and strong-willed people; and by this fact, he accuses the externalist of ascribing derivative desire to good people's motivation. In other words, the externalist explains the reliable connection by appealing to factors that are not directly derived from agents' moral judgements. Precisely, according to him, the most his opponent can offer is a reliable connection that is based on whatever the agents regarded as the right thing to be done. Note that the cause of disagreement is not necessarily desire, because Smith believes that desire is a crucial motivational factor. His worry hangs on how desires figure in agents' motivation. According to him, at least in a good and strong-willed person, the desire to do the right thing is *non-derivative*. On this reading, Pete's motivation to oppose the government's interference is non-derivatively, that is, not by some sort of general desire to do whatever that is right. Smith thinks that the externalist construal of good people's motivation is problematic, hence, against our ordinary understanding:

For commonsense tells us that if good people judge it right to be honest, or right to care for their children and friends and fellows, or right for people to get what they deserve, then they care non-derivatively about these things. Good people care non-derivatively about honesty, the weal and woe of their children and friends, the well-being of their fellows, people getting what they deserve, justice, equality, the like, and not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re.* Indeed, commonsense tells us that being so motivated is a fetish or moral vice, not the one and only moral virtue. (Smith 1994, pp. 74 – 75)

Notice that any agent that fails to satisfy Smith's purported construal is not a good person. By implication, Smith is accusing the externalist of making good and strong-willed people appear like 'bad' persons. It is to this end that he advises us to adopt internalism because it is not susceptible to such a problem.

Let us unpack Smith's claim. First, to what extent does his purported construal of commonsense capture folk understanding of good people? Do we consider people who do things because 'it is the right to do' as bad people? Should we regard people like Maximillian Kolbe, who desired to die for a stranger (given his religious belief, *it is the right thing to do*) as a bad person? Definitely, on Smith's construal, he is not a good person, because good people care about those things non-derivatively. What about the millions of 'ordinary people' who sacrifice their lives for the sake of their families, communities, nations? Does the thought that it is right to help millions of women, men and children fleeing their home because of wars or hunger make, them bad people? Alternatively, do we not still celebrate people who thought it is right to stop slave trades, killing of twins? Is it against commonsense to call them good people? I doubt very much if Smith's purported commonsense of good people is actually a commonsense. In other words, his enthroned commonsense construal is, in fact, an idea he borrowed from Bernard Williams.

In *Persons, Character, and Morality*, Williams (1981) discusses the case of a man who, while confronted with the situation of choosing between saving the life of a stranger or that of his wife, chooses to save that of his wife. He argues that the man saved his

wife's life by the mere fact that the person drowning is his wife. He did not save her because it is right or permissible to save her in this situation, but simply by the single thought that the person drowning is his wife. The argument is that the man's concern to save his wife is direct, for if it were not so, he would be entertaining 'one thought too many'. Adopting this idea, Smith thinks that good people act morally for a direct concern:

> Just as it is constitutive of being a good lover that you have direct concern for the person you love, so it is constitutive of being a morally good person that have direct concern for that you think is right. (Smith 1994, p. 76)

In essence, Smith accuses the externalist of making moral agents engage in 'one thought too many'. On this reading, Smith would argue that motivation of people like Maximilian Kolbe did not issue from the direct concern; hence, they are not good people. However, his argument is not conclusive because it is also part of our commonsense experience to regard people such as M. Kolbe and people mentioned above as good people. By claiming that his construal of good person is final, Smith would only be begging the question against our ordinary understanding of good people.

Furthermore, Smith's accusation that externalism makes good people engage in 'one thought too many' is false, because Williams' construal is wanting. Contrary to the claim that the man saved his wife by direct concern, Harry Frankfurt (2004) argues that the mere thought that the woman drowning is his wife, makes the man engage in 'one thought too many' as well. For Frankfurt,

> Surely the normal thing is that he sees what's happening in the water, and he jumps in to save his wife. Without thinking at all. In the circumstances that the example describes, any thought whatever is one thought too many. (Frankfurt 2004, p. 36)

On a similar note, Miller argues that Smith's 'one thought too many' strategy against externalism undermines his internalism as well. He writes,

> According to Smith's internalist, if George judges that honesty is right, it follows, as a matter of conceptual necessity, that George will be motivated to be honest. Now note that on this account George's motivation to be honest is still derived: not, to be sure, from his belief that it would be right to be honest in conjunction

134

with an over-arching non-derivative desire to do the right thing, but from the belief that it would be right to be honest itself. (Miller 2003, p. 225)

Given the 'one-thought-too-many principle', agents both on the externalist and internalist explanations seem to be bad people because their desires to do the right thing are derivative. Smith's attack on externalism not only fails, but it backfires on his internalist project. Smith conceded to this objection, but he did not rest his case. (Even if Smith's non-derivative arguments were to be true, it would as well be false to claim that externalism lacks an account of motivation based on non-derivative).[77]

Smith readjusts his position from *non-derivative desire* versus *derivative desire* to *non-instrumental desire* versus *instrumental desire.* The upshot of this readjustment is Miller's objection. In a reply, Smith argues that on the internalist explanation, by judging that it is right to φ in C, agents do not do that 'in virtue of any further right-making feature of φ-ing. They judge it fundamentally right to φ in C, simply as such' (Smith 1996, p. 179). They do not do the right thing for the sake of realizing other goals. Hence, unlike the instrumental conception, an agent's desire to do the right thing is non-instrumental. If it is true that externalist agent's desire is both derivative and instrumental, then it could be argued that the agent's moral reason is both responsive to whatever an agent thinks is right and whatever he thinks would satisfy his contingent desire. The internalist seems not to be exposed to this threat, for according to him, although an agent's desire might be

---

[77] Possibly, an agent's desire for motivation "stems from the fortuitous connection of her new moral judgement and some pre-existing desire. And if a conscientious or virtuous person can have a battery of non-derivative desires, each of whose content is quite broad, then the dismissal is probably too quick. In the example Smith gives to illustrate the externalist's need to assign just the motive of duty as the explanation for the reliable connection between judgement and motivation, a good and strong-willed person who has supported the libertarian party, but is now convinced that the social democrats are better deserving of support, will be motivated to support the social democrats and abandon support of the libertarians. It seems, though, that the externalist can explain this by citing a quite general desire to support just political institutions, or more egalitarian political parties. This general desire, accompanied by the new political/moral belief about the social democrats, generates a derived desire that is sufficient to explain the change in motivation. And this sort of explanation would work for all cases in which the change in moral belief leaves the propriety of one's non-derivative desires intact. For instance, one might have fundamental, non-instrumental desires to see one's family flourish, to see justice done, to promote the welfare of the worst off in society, to work towards gender equality, etc. These are all quite general desires that will survive most changes of moral belief. These desires, and others like them, will be able to explain why good agents gain new desires when acquiring new moral beliefs" (Shafer-Landau 1998, p.365).

derivative as we saw above, it is non-instrumental. In an example, Smith explains how this idea of non-instrumental desire figures in an agent's moral judgement and motivation.

> Suppose that the person is a utilitarian and that he judges it right to maximize happiness and minimize suffering. The internalist is bound to insist that the motivation this agent must have, if he makes this judgement and is not weak-willed, is non-instrumental in character: that is, the utilitarian's motivation to maximize happiness and minimize suffering is not dependent on the fact that its satisfaction contributes towards the satisfaction of some further motivation that the agent has. The maximization of happiness and minimization of suffering is the ultimate consideration that moves the agent. This is a direct corollary of the fact that the judgement concerns fundamental rightness, according to internalist. (Smith 1996, p. 179)

Suppose the leader of the GHESx Party believes that it is right to take care of unaccompanied asylum seekers (that is, minors). He desires to minimize their hardship. Given Smith's proposal, the internalist must insist that this desire is borne out of the fact that it is fundamentally right to protect this group of asylum seekers. However, if the internalist is to be taken seriously, he has to tell us how the party is to realize this desire. Suppose the leader decides to provide them with a good education. Notice that the internalist construal admits means-end belief; hence, in order to minimize their hardship, these minors should, for example, be sent to school.[78] In other words, the leader does not just desire to minimize their hardship, he desires that they should be sent to school. Thus, his desire is instrumental. If practical rationality is, at all, relevant to morality, then it is difficult to see how the conception of non-instrumental desire is supposed to function in morality. For the sake of argument, let us assume that the desire to send these minors to school is non-instrumentally derived from the fundamental right. The leader would still be faced with a subset of instrumental desires such as: What type of education best suits the needs of these asylum seekers given their religious, cultural and language backgrounds. He cannot just send them off to 'one size fits all' school? Assuming he finds

---

[78] Since Smith believes that Hume's picture of human psychology is true, I do not think that Smith would reject this means-end belief.

the best school, it is also possible that he would be faced with further subsets of means-end situations. It is hard for Smith's agents to escape the scope of instrumental desire based on this construal. Thus, we can as well accuse him of making agents' desire both derivative and instrumental.

However, Smith might contend that such additional features do not fit into the picture of a morally perfect person. In essence, he could object that such instrumental conception is the stock in trade of externalism because the internalist conceives a morally perfect person to be motivated solely by the feature of his judgement that it is right to do something. However, his morally perfect agent meets another objection, namely: How to handle a psychology that solely builds its moral motivation on moral judgement. How does Smith's understanding of a morally perfect person fit into ordinary conception of moral perfection? Apart from shying away from giving us a full characterization of what constitutes a morally perfect person (the much he did was to assume him and claim that he is solely moved by his moral judgement), Smith's conception of a morally perfect person is problematic because it mutes other psychological factors that link up to constitute moral motivation in ordinary people.[79] Miller was right when he pointed out the woman's complaint in Williams' story

> [I]s that there is some aspect of the husband's psychology, in addition to the belief that the woman drowning is his wife, to which his motivation to jump in and save her is sensitive, where, crucially, it does not matter whether or not the other aspect of the husband's psychology is a desire or whatever. (Miller 2003, p. 227)

Miller's contention is (assuming Smith's morally perfect person derives his motivation non-instrumentally) that the man's belief, namely 'it is right to do something', is still sensitive to some features of his psychology. Therefore, "as in the case of the husband with 'one thought too many', this is enough to merit that charge that the psychology of

---

[79] It goes without saying that Smith's conception of morally perfect persons or good and strong-willed persons faces a similar accusation that Shafer-Landau levelled on McDowell, namely that such a construal restricts moral motivation to only few population people. "But even if […] the picture of the virtuous agent's judgments is correct, and even if it is plausible that such an agent would thereby be motivated to act in accordance with her judgments, we have a defense of motivational judgement internalism only as it applies to virtuous agents. That is a quite serious restriction of internalism's scope" (Shafer-Landau 2000, p. 280). Such a claim makes the idea of widespread motivation objectionable.

the morally perfect person has been misinterpreted: in the morally perfect person, the motivation [that it is right] would be sensitive to no such aspect" (Miller 2003, p. 227).

5.7 Conclusion

In this chapter, we have seen that Smith's defense of internalism both from rationalism and practicality requirement did not succeed. His claim that moral requirements apply to rational agents as such is false because there are cases where being rational and being moral come apart. Moreover, even if it were the case that moral requirement binds moral agents, it does so under agents' construal of moral reason (where the reason is not just Smith's account of normative reason). Against the practicality requirement, we have seen that Smith's construal of a good and strong-willed person is not conclusive. Above all, his attack on externalism based on derivative and instrumental desires is not sufficient to convince us to adopt internalism.

Chapter 6

OBJECTIONS AGAINST EXTERNALISM AND REPLIES

> Internalism overstates the connection
>
> between morality and motivation…
>
> [I]t does not take the amoralist's challenge seriously enough.
>
> ***David O. Brink, Externalist Moral Realism, 1986***

6.0 Introduction

In the last two chapters, we considered both the robust and conditionalized versions of realist internalism; and we found their claims to be implausible. In this chapter, I consider four main objections against externalism and attempt to dismiss them.

6.1 Objection from Moral Intuition

Objection: Our intuition about morality is that motivation directly follows from moral judgement. Therefore, the externalist position flies in the face of our ordinary moral experience.

Reply: I do not think any serious externalist position would deny that moral judgements have practical influence on our everyday moral experiences. After all, the task before the externalists and internalists is that of explaining the ubiquitous connection between moral judgement and motivation. Nevertheless, we must not be internalists to explain this intuition. For example, we can claim that motivation follows indirectly from moral judgement; and this is not inconsistent with our moral experiences. However, it is too early to draw a conclusion because there is more to the intuition as well as conceptual truth on which the realist internalist grounds his argument.

Intuitions are pre-theoretical understandings of truth. They are "simply opinions; our philosophical theories are the same. Some are commonsensical, some are sophisticated; some are particular, some general; some are more firmly held, some less. But they are all opinions…." (Lewis 1983, p. x) If only roughly, intuition is supposed to be a largely shared understanding of the truth of something. Our intuition about something, for example, that *p* is true, is revised in a case of theoretical justification against *p*. That said, the internalist employs our intuition about morality to explain the conceptual truth of motivation. When we say that *p* is a conceptual truth, it is supposed to be a part of our intuition that *p* is true because of what *p* means. Hence, the internalist argues that it is a conceptual truth, that is, part of our intuitive understanding of moral judgement that it motivates efficaciously. The essential factor driving conceptual or analytic truth is the *apriority* claim. That is, the claim that direct flow of motivation from moral judgement is not dependent on experiential evidence (hence, *a posteriori* claim). To the internalist mind, motivation is a sign that an agent (without psychological or cognitive impairments) has made a genuine and sincerely first-person moral judgement; and if motivation is missing, he claims that we cannot assign moral judgement to such an agent. Notice that moral judgement seems to acquire a 'new' definition. To illustrate:

D1: Genuine and sincere moral judgement is true if and only if it motivates.

This definition seems to work at least superficially. However, unlike in the case of *p* (*p* is a stipulation, such that *p* is true if it satisfies certain conditions or features assigned to it), motivation occurs within human psychology; and it has basic features that cause our experience of it, hence our moral experience of a motivated agent. Some aspects of these features or conditions, in the case of motivation, are satisfied experientially and independent of stipulated conventional features, unlike in the case of p. Notice that if our moral experience of motivation is caused by these basic features, then our knowledge thereof depends on some sort of experiential evidence. Suppose that the internalist is correct – that motivation is the sign that an agent made a genuine and sincere moral judgement (this is supposed to be our meaning of moral judgement). Suppose we observe

140

that being motivated entails, among other things, displaying x and y features. Hence, to be motivated is to have made a genuine and sincere first-person moral judgement – this is a conceptual truth and not an experiential truth. However, we can as well say that an agent is motivated when we observe him doing acts that have x and y features – this is experiential truth and not conceptual truth (I don't think the internalist would have problem admitting this since he accepts that being motivated directly by moral judgement is part of moral experience). Given that it is a conceptual fact that motivation follows from moral judgement, Rudder Baker (2001) rightly argues that the conceptual analyst claims that the features of x and y follow *a priori,* such that we can conclude that x and y follow intrinsically from moral judgement of x and y.

Nevertheless, the features of x and y are not known *a priori*, but experientially from motivated agents. Adopting Rudder Baker's move, we can show that the conceptual claim of the internalist, if not false, is insufficient to explain the experience of moral motivation.

D2: It is conceptually incorrect to attribute a genuine and sincere moral judgement to an agent if his actions do not possess x and y features (i.e. being morally motivated).

It is part of our *meaning* of moral judgement to capture it thus. According to Rudder Baker, the conceptual analyst holds the following:

(1) True meaning statements are available for use in philosophical arguments, and
(2) Meaning statements are justifiable *a priori* (Rudder Baker 2001, p. 384).

The internalist thinks that the meaning of the concept of moral judgement is sufficient for our explanation of (psychological experience of) motivation. However, Rudder Baker rightly argues that we have to *suppose* certain features in order to establish an *a priori* claim. In our case, we supposed x and y such that when we see an agent exhibiting them, we know that his motivation directly follows from his moral judgements with the content

of x and y.[80] But suppositions are insufficient to establish *a priori* truth. Rudder Baker argues that, "we must *suppose* that we have a true meaning statement (that is to be regarded as a priori); no one claims that we actually have one" (Rudder Baker 2001, p. 384). Given that we have to suppose them to make our concept of moral judgement (generating motivation efficaciously) an *a priori* claim, namely, conceptually true; it is inconclusive to claim that a belief based on such a supposition is sufficient to explain human motivational experience. Hence, assumption (D1) seems to be implausible.

Assumption (D2) is related to (D1) in the following way: Even if it is the case that it is conceptually true that motivation follows from moral judgement, it would still be insufficient to explain human motivational experience only on the basis of analytic truth. In (D2) the internalist claims that we can justify the claim that "it is conceptually inconsistent to assign a genuine and sincere moral judgement to an agent in the absence of motivation" without relying or employing any experience, hence *a posteriori* means. However, unlike in the case of p, human motivation is not a pure analytic experience. To claim otherwise is to replace human psychology of motivation with a mechanics of analyticity. Or better put, to claim that moral motivation is explainable essentially with conceptual resources is to reduce such experiences to what I refer to it as *conceptual psychologism*. As we noted above, motivation generally occurs within a complex human psychology; and there are certain basic features of motivation (occurring in agents) that cause our experience of it. Conceptual resources, in the absence of experiential resources, do not exhaust moral motivation. The claim that motivation necessarily follows from sincere first-person moral judgement alone flies in the face of human motivation generally. This is because motivational experience can occur in agents without (or in the absence of) conceptual resources. Such agents are motivated, for example, naturally in an intentional way.[81] Given that moral motivation is a species of human motivation; and that

---

[80] Nevertheless, a caveat is in order here. It would be false to claim that there are no cases (such as 'A sister is a female sibling' (Rudder Baker's example), a triangle has three angles) where the meaning statements are available for philosophical argumentative usage. Nevertheless, such few cases are insufficient to build the foundation of any meaningful theoretical inquiry.

[81] Roughly, the conceptual claim does not exhaust motivation because motivation occurs in primates without sophisticated linguistic apparatus like us. Just as we do not necessarily need language for perception, it also appears to apply to motivation in general. Assuming there are cases where motivation occurs in non-moral contexts without formation of judgement, it would be counterintuitive to claim that

the latter can occur in agents without linguistic apparatus, it is a bit odd to claim that moral motivation is only possible because of the agents' utterance of first-person sincere moral judgements. In other words, even if we consider linguistic apparatus as a crucial aspect of (moral) agentic capabilities, it would be false to claim that it presupposes internalism, because we can analyze the connection in externalist terms as well. However, analyzing moral motivation solely in terms of linguistic resources seems problematic, because of the vital roles other psychological factors. Precisely, the *possibility* of moral motivation may very well rest on linguistic (and other rational) capacities, but the question of whether an agent is *actually* motivated to act according to his or her moral judgment is still open and may rest on additional psychological factors. Thus, it is still difficult to see how the internalist can justify motivational experience solely with his conceptual resources. Externalism, on the other hand, seems better positioned to locate the possible factors that underpin our widespread experience of moral motivation in agents.

Let us for the time being suppose that internalism is supported by folk intuition, that is, that people think that moral judgement necessarily generates motivation. On this view, the internalist claims that *people* actually see the cases of moral indifference as flying in the face of how we understand morality. Against this appeal of folk intuition, there is current empirical research that shows that the externalist cases of amoralism and moral indifference are not at odds with how people think of the connection between moral judgement and motivation. Consider here as studies in point Strandberg and Björklund (2014), who report that folk intuition does not largely support internalist claims. In six scenarios, the study tests the folk intuition about the claims of generic internalism, conditional internalism and communal internalism (the view that the necessary connection between moral judgement and motivation occurs at the community level and not on the individual level). The result shows that the participants responded in a way that "provides reasons to be skeptical of the evidential basis" of these versions of internalism

---

moral judgement alone is the necessary and sufficient requirement for moral motivation (given that moral motivation is species of general motivation). The claim that motivation seems to occur in closer primates without linguistic apparatus for making judgement, points to a crude fact that what drives motivation must be outside of moral judgement. In other words, moral judgements might be seen as sufficient, but not necessary for motivation.

(Strandberg and Björklund 2014, p.16). In three experiments, Leben and Wilckens (2014) reveal that the internalist purported common intuition is not widespread. Further, in a neuroscientific and psychological study, Roskies (2003, 2006) documents that this intuition is undermined by acquired sociopaths; and Nichols (2002, 2004) argues that the psychopath's case[82] threatens conceptual rationalism, the sort defended by Smith (1994).

Of course, there are counter-studies from internalists such as Cholbi (2006, 2011); Kennett (2006); Maibom (2005); Kohlberg (1981).[83] That said, the goal here is not to claim that externalism has won the intuition war. Rather, we want to show that the internalist reliance on folk intuition to confirm the connection between moral judgement and motivation is shaky, because the tug of intuition is far from being conclusive. These studies seem to show that folk intuition is commonly shared by externalism as well as internalism. Given that the internalist cannot monopolize our common intuition, the externalist can further propose explanations that account for a reliable and ubiquitous connection between moral judgement and motivation: An account that will, among other things, explain the motivation of virtuous persons, 'ordinary' moral agents, as well as the phenomenon of moral indifference.

6.2 Objection from Action-Guiding and Motive-Giving

Objection: Moral judgements are both *action-guiding* and *motive-giving*. Hence, the externalist case of moral indifference is odd, because to make sincere moral judgement is not just to accept moral reason, but also to be motivated by it.

---

[82] For other studies in psychopathy, depression vis-à-vis moral motivation see Barry, (2013): *Psychopathy and Moral Judgement*; Cima M1, Tonnaer F, Hauser (2010): *Psychopaths know right from wrong but don't care*; Blair, Jones, Clark and Smith, (1997): *The Psychopathic Individual: A Lack of Responsiveness to Distress Cues?*

[83] Since Kohlberg's work, it is observed that psychologists researching on morality appear to have accepted the truth of internalism as a dogma; and this stance is explainable based on Kohlbergian Kantian and Socratic (knowing good is doing it) assumptions on the one hand; and the assumptions of Aristotle's virtue ethics on the other hand. Against this stance, Kristjánsson argues that academic moral psychology is not condemned to internalism. In fact, not only is externalism possible on Aristotelian assumptions, but also such an externalist position enables moral psychology to "advance truths about the good, virtuous life while remaining a science" (Kristjánsson 2013, p. 422).

Reply: The internalist is right in claiming that moral judgements guide and motivate us morally. But it is false to assume that accepting the action-guiding force of moral judgement is to be motivated necessarily. The internalist makes a double-necessity claim on moral practicality. On this view, moral judgement is not only motivationally efficacious, but also, it necessarily guides actions. By the very fact that moral judgement provides us with requirements, it intrinsically provides reason for action. For him, it is odd not to be motivated after accepting its moral requirement. This implies that moral judgement has both intrinsic motivational and prescriptive efficacies. Notice that the internalist strategy is to secure the binding force – the authority of moral judgement. It looks embarrassing, on the internalist framework, that being moral, in the absence of psychological impairments or overriding motives, is not being motivated to act accordingly. The embarrassment is rooted in the following: It is thought that such a person threatens morality. However, this fear is unreasonable, because it is one thing to be morally obliged and another thing to be motivated. What makes the moral demands binding is the truth determined by moral facts and not the agent's desire to be motivated by such truths. Zangwill rightly writes that,

> What makes our moral judgements true or correct is one thing, and what motivates us is another. There is no reason why one should not hold that the moral truth is independent of our desires, while moral motivation is dependent on them. (Zangwill 2003, p. 146)

The internalist problem is that he fails to see the line separating truth of moral judgement from the fact that our motivation contingently depends on such truth. By collapsing the wall, the internalist forcefully eliminate the agent's desires from the dynamics of moral motivation. He assumes that to accept the truth of morality is the same thing as to be motivated. However, Brink argues that

> [It] seems just false to the psychological facts. Though indifference to what are regarded as moral considerations may be fairly rare, it does seem to exist. Some people (e.g. certain sociopaths) do not care about moral considerations. (Brink, 1986, 29)

145

This shows us that [1] recognizing or accepting truth of morality does not necessarily guarantee actual moral motivation; and [2] motivational failure does not threaten the moral demands or bindingness on moral agents, regardless of their motivational states. In other words, there is nothing embarrassing about moral judgement and motivation coming apart. This is because an agent might accept what morality requires without being motivated. One might understand the truth of a moral requirement and yet fail to care about it. However, this stance does not imply that *moral bindingness* does not apply to such agents. Moral requirements bind both the motivated and unmotivated, but they might not be necessarily motivated. Shafer-Landau rightly captures this concern thus: "We do believe that moral obligations may apply even (and especially) to those whose attitudes imply their rejection, and this is something that internalists have had difficulty accounting for" (Shafer-Landau 2000, p. 278).

Another way of making this point is by asking whether moral normativity entails moral motivation. We might dare to answer affirmatively, but such an answer sticks only in cases where motivation occurred. However, we have to quickly add that the motivation in question must not be understood in the internalist term. In other words, moral normativity entails motivation in cases where agent's moral judgements link with other distinct desires. Once again, normative character of morality essentially applies even to morally indifferent agents, but it does not follow that the truth of moral normativity is motivationally efficacious. In sum, we can accept the connection between moral normativity and motivation provided that there is a distinct desire that makes moral motivation a contingent fact. Hence, in the absence of such desires agents can be morally indifferent to the requirements of their sincere moral judgements.

6.3 Objection from Off-Colour Moral Terms Usage

Objection: An agent who upon making a moral judgement remains unmotivated must be using the moral terms in an 'inverted commas' (Hare 1952, pp. 124 – 126); or 'off-colour' manner (McNaughton 1988, p. 136) or like a '(born) blind' person who makes judgement about colours with the help of supporting facilities (Smith 1994, pp. 69 – 71).

Reply: Generally, internalists employ this strategy to diffuse the externalist counterexamples, for example, the cases of amoralism, moral cynicism, moral indifference, psychopathy etc. Our concern here is on realist internalism. Precisely, when a realist internalist says that agents (such as Olivia) use moral terms in inverted comma's, it appears harmless at face value. However, he unwarily uses his opponent's weapon to fight a further opponent. Strictly, 'inverted comma's strategy is not available to the realist internalist because Hare's construal of moral judgement is radically different from that of the realist. Being an emotivist, Hare essentially sees moral judgement as a prescription (hence, prescriptivism) that does not describe natural facts – non-descriptivism (Hare 1952, pp. 2, 79 – 93). In other words, using moral terms properly on his analysis differs from what the realist calls moral terms figuring in moral judgements.

> If moral judgements are individuated as the noncognitivist says, then judgements employing different moral predicates must be accompanied by different attitudes. Moral judgements are distinguished not by the different properties that are being referred to in the different judgements, but instead by the different associated attitudes that receive expression in moral judgements. (Shafer-Landau 2000, p. 279)

In other words, while internalists (non-cognitivist and realist) have a common opponent – externalism – realist internalism is to be wary in employing strategies that appear to be common to internalism.

However, the internalist realist can sustain his argument by dropping Hare's 'inverted commas' usage. He can still reject the claim that the unmotivated agents in question make genuine moral judgements in the first place. On this view, we cannot assign genuine moral judgement to morally indifferent persons. To the internalist mind, they (i.e. indifferent persons) do not make use of moral terms as 'normal' users do. Notice that this argument rests on the confidence that genuine moral beliefs are on their own motivationally efficacious. However, I doubt very much whether this claim is true. It is not the failure of making genuine moral judgement that leaves agents unmotivated; agents, in fact, are unmotivated because their genuine moral judgement is inert. In other

words, moral judgements are genuine beliefs; and an additional in-built motivational character is unwanted.

Both the realist internalists and externalist believe that moral judgements are genuine beliefs, but the internalist thinks that in addition to being genuine beliefs, they are motivationally efficacious. Moral beliefs have similar ontological and semantic characters with ordinary beliefs. Both beliefs have similar propositional attitudes, representational, assertoric and truth aptness. It is disturbing that the internalist assigns motivational force to just moral belief and not to the other genuine beliefs. After all in chapter 4, Bromwich pointed out that the realist internalist needs to argue that all beliefs, not just moral beliefs, motivate *simpliciter*. However, the internalist claim is strong because he is not claiming moral beliefs at some point acquire motivational force, but the force is an in-built feature of moral beliefs. If ordinary beliefs and moral beliefs are genuine belief, hence *genuine geese:* Why is it just good for one goose? Why is it not also, at least, good for the gander (if we suppose that the ordinary belief is a gander)?

Furthermore, the internalist's usual move is to claim that moral beliefs, unlike ordinary beliefs have special character, and this character is both *moral* and *motivational.* Consider the following:

> Supposing that I could play tennis and one of you saw me playing and said 'Well, you play pretty badly' and suppose I answered 'I know, I'm playing pretty badly but I don't want to play any better,' all the other man could say would be 'Ah, then that's all right.' But suppose I had told one of you a preposterous lie and he came up to me and said, 'You're behaving like a beast' and then I were to say 'I know I behave badly, but then I don't want to behave any better,' could he then say 'Ah, then that's all right'? Certainly not; he would say 'Well, you ought to want to behave better'. (Wittgenstein 1965, p. 5)

Wittgenstein claims that there is something special about saying:

(Ai) "You're behaving like a beast"

(Aii) "You *ought to* want to behave better".

148

This special character, so to say, imposes moral categoricity on us, hence being moral is being motivated, at least to some extent. However, are (Ai) and (Aii) ontologically and semantically different from?

(Bi) "You play pretty badly"

(Bii) "You *ought to* want to play better"

We should not forget that moral belief participates, roughly, in the genuineness of ordinary beliefs. Ordinary beliefs are not credible because they are like moral beliefs, but the other way round. That said, by claiming that moral beliefs are both categorical and motivational, the internalist is attributing a foreign character to moral beliefs. Assuming this special character exists, it adds nothing to the credibility of moral beliefs (that is, it adds nothing to belief-ness of moral belief). Rather, the additional character not only makes moral beliefs dubious, but also motivationally suspicious, because it is that purported character that is supposed to explain the psychology of moral motivation.

By saying, "You play pretty badly", the speaker is passing a value-factual judgement. This is not a normatively lame statement. Suppose the player is aiming for a trophy, in other words, he ought to want to play better. He cannot just say: 'I know, I'm playing pretty badly but I don't want to play any better'. Suppose the other man is his coach, it would be stupid of him to reply: 'Ah, then that's all right.' Notice that normatively (in addition to the ontological and semantic status), there is no difference between tennis player and the 'moral player' apart from the 'context of the game'. Foot has argued in similar direction, the 'ought of morality' and the 'ought of etiquette' or 'club rules' are no different. In other words, 'moral ought' is no more categorically special than the 'ought of etiquette' (Foot 1972).

More so, even at aiming for a trophy, the tennis player is not absolutely bound to play any better, if he does not care about the trophy. He can say, "I *know* that winning this trophy would make me the best tennis player" but "I don't care about it". So also, there is nothing special about saying: "I *know* I can behave better" but "I don't care about it". Both players *know* what is required in their respective games – game requirements. Notice that morally indifferent persons, contrary to not making genuine moral

judgements, can decide not to care about the truth of morality, even when they are bound by it (as we mentioned above). As it is, there is no special force hidden in moral judgement that necessarily morally binds and motivates. Agents' moral judgements are genuine beliefs, but it is one thing to judge and another thing to care: To care is an emotional capacity, and they (that is, care about morality, moral-self and moral desire) work in tandem to provide us with motives for acting on genuine moral judgements (as we shall see in the next chapter).

An alternative way of making this point is thus: The externalist believes, so also the realist internalist, that moral beliefs are objective and mind-independent truth-values. To judge is not to care or express attitudes, rather to represent things as they are in the world. On the realist construal, to judge is to represent. Thus, by claiming that motivation is internal or in-built in moral beliefs, the realist internalist is assuming that to judge is 'to express care or concern'.[84] However, such resources seem to be open only to the non-cognitivist, because according to them, moral judgement are attitudes, cares and emotional expressions, and all these expressions are necessarily motivating even if only weakly. Hence, to judge is to express concerns or desires. The realist explicitly rejects this view, but tacitly the realist internalist confirms the non-cognitivist intuition. This informs why the realist has to be wary about espousing internalism: "[B]elieving in motivational internalism gives one a certain impetus to be a moral anti-realist, as anti-realism and internalism seem to form a coherent and convenient interlocking set" (Kristjánsson 2013, p. 426).

Of course, the realist internalist is free to deny any surreptitious attraction to noncognitivist motivational commitments. However, he still bears the burden of explaining the constituting factors of moral actions (as a species of intentional action); that is, the roles of belief and desire in moral actions. Again, given the cognitive content of moral judgement, the realist internalist has to explain how desire figures in moral judgement to crank out motivation. Well, at this juncture, he has various options: It might

---

[84] Given that externalism holds that motivation is external to moral judgement, an agent's care about morality is located in a distinct capacity and not in moral judgements. So, the claim that moral judgement represents is compatible with the idea that an agent can care more or less about morality. This idea is also available to non-cognitivism because to judge is the same thing as to express concerns, inclinations, or desires.

be the case that moral judgement has both belief and desire content, thereby forming some sort of unitary psychological state. However, collapsing the two states into one state does not help that much, because this unitary psychological state ends up looking like Orwellian pigs in the *Animal Farm* (that is, identity crisis). Further, he might insist that desire is only an incident state. In that case, he would be presenting to us, as McNaughton rightly notes, with:

> The chilling, and implausible, picture of the virtuous person as someone completely cold and uncaring who simply sees what is right and does it without emotion or concern. Feeling or passion would have no place, except as incidental accompaniments to the entirely unemotional process of recognizing the moral facts. (McNaughton 1988, p. 113)

There is another open move, namely, a realist internalist can as well be a Humean about motivation (Smith defended this position in the *Moral Problem* 1994). This option amounts to admitting that moral belief conjoins with desire to motivate. One of the implications of this position is admitting that moral beliefs and desires are distinct psychological states following different directions of fit. It goes without saying that realist internalist Humeanism comes with a cost. This is because it is difficult to see how one is supposed to handle Humean psychology that is instrumental on the one hand, and holding a realist internalist creed, on the other hand. The task before the realist internalist about Humean motivation is that of explaining how moral beliefs relate to desires and at the same time motivate agents non-instrumentally and non-derivatively.[85] The most we could get from this position is from Smith's claim that moral beliefs intrinsically link up with desires to produce motivation. Nevertheless, even at that, it is difficult explaining how these distinct psychological states figure in internally to generate motivation. In sum, externalism does not shoulder these burdens, because desires are separate states, and

---

[85] If Hume is correct about the distinct characters of beliefs and desires; and about the claim that beliefs are motivationally inert in the absence of accompanying desires, then it seems that a realist internalist espousing Humeanism about motivation is absurd. This is because he accepts Humean motivational resources and at the same time insists that moral beliefs on their own are motivationally efficacious. Externalism does not suffer this threat, because it is compatible with Humean motivational psychology.

external to moral judgement. It is these distinct desires that provide agents with the motives for acting on moral judgements.

6.4 Objection from Amoralism

Objection: The amoralist must be strikingly different from others, because it is not possible that someone could be in the same cognitive state as a 'normal' moral agent and fail to be motivated by sincere moral judgement (McNaughton 1988, p. 137, 139).

Reply: On the contrary, it is possible to be in the same cognitive state and yet not be moved by the demands of moral judgement. Amoralism continues to attract wide attention especially in the motivational debate; and sometime characterizing amoralism appears to be difficult (as we mentioned in the previous chapter). We shall attempt to define amoralism or amoralist behaviour, first, by saying who an amoralist is not. Going by its etymology, 'a'- 'moral'- 'ism' paints the idea of someone who is 'against' morality. Strictly speaking, the amoralist is not like a bad or immoral person, who accepts morality but occasionally he chooses to go against the authority or demands of morality, due to error in his moral judgement or competing personal interests. Bad persons are "people who have some commitment to morality, and so see moral considerations as providing some reason for acting, but who allow other commitments to play a larger role in their lives than they should" (McNaughton 1988, p. 145, see also Kristjánsson 2013, p. 429). The amoralist is unlike an evil person, who accepts morality, and sees the good that is required to be done, but he chooses the opposite, namely, he does evil for its own sake (Kristjánsson 2013, p. 429; Benn 1985). He declares, like Milton's Satan, "evil, be thou my good!" Finally, the amoralist is not like an akratic person, who accepts, sometimes even full demands of morality, but he is too weak to act according to his 'best' practical judgement. He suffers some sort of "internal conflict" like Paul: "For I do not do the good I want to do. Instead, I keep on doing what is wrong I do not want to do" (Romans 7:19).

The amoralist doubts and even rejects reasons to be motivated morally. He sees no reason to follow the demands or authority of his moral judgement. He does not necessarily need to reject morality due to his egoist purposes (that is, rejecting morality

152

because the requirement in question does not satisfy his self-interests). It might be due to his suspicion regarding the underlying character of the purported rationality of a given moral standard or institution. McNaughton notes that some non-moral ideals might totally block someone from giving moral considerations any weight in his life. "Such an attitude is perhaps exemplified by those who place the value of art above that of human life, like the aesthete who did not care that thousands died to build the pyramids because the result was so glorious" (McNaughton 1988, p.135). Of course, these factors are far from exhausting amoralist grounds for rejecting the authority of morality.

More so, amoralist cases come in various flavours and degrees. On the one hand, there are cases of amoralism that are linked to psychological derangements. We may classify sociopathic, psychopathic and antisocial behaviours under this group. If the internalist charge sticks, then this class of amoralists is not in the same cognitive state with 'normal' moral agents. However, it is not clear what the internalist means by cognitive state. But let us assume that he means the state that underlies first-person moral judgements. Nonetheless, for example, if it is the case that the psychopath's cognitive faculties have little or no impairments (Prichard 1974, p.15); and he uses moral terms as well as form moral judgements in no off-colour sense (Roskie 2003; Damasio 1995; Damasio et al. 1990; Hare 1993), then it is odd to suppose that they fail to fulfil the cognitive state requirement. It might not be odd if we assume that such cognitive state is accompanied, if only incidentally, by some other factors such as emotions or conscience. In that case, we can rightly accuse amoralists of not sharing the same dispositional states with 'normal' moral agents. Nevertheless, this move is incompatible with the internalist thesis, because he claims that it is only moral judgement that motivates. If it is true that the amoralist motivational failure is not due to intellectual or cognitive errors, then the internalist objection is discharged. In fact, it is argued that the internal causes of antisocial behaviours are closely tied to deficiencies in emotions such as shame, guilt, fear of punishment, empathy (Damasio and Damasio 2000; *Diagnostic and Statistical Manual of Mental Disorders* IV), sometimes conscience and concerns for other persons are taken to be the cause (Lykken 1995, p. 115, Hare 1993, pp. 75 – 76).

On the other hand, there are cases of amoralism that are not due to psychological derangements – cognitive and conative impairments. This group of amoralists is aware of the authority of morality, and accepts the demands of moral judgements. But they *see* no reason to act according to them. People such as Olivia and Zangwill's mercenary belong to this group of amoralists; and it is this group that Foot had in mind when she wrote, "One *can* be indifferent to morality without error"; and reject "morality because he sees no reason to obey its rules" (Foot, 1978, pp. xiv, 161). This type of amoralism comes in degrees.[86] Brink observes that the weak amoralist sees reasons, but they are not sufficient to motivate her, whereas the strong amoralist does not see reason at all to be moral (Brink 1989, p. 60). The strong amoralist rejects moral institutions or standards. McNaughton thinks that the internalist might want to adopt this version of amoralism, because it enables him to dismiss the externalist claim that the amoralist accepts the existence of a moral requirement, but he sees no reason to follow it. In other words, the strong version paints the amoralist as someone who does not share moral categories and frameworks with normal moral agents (McNaughton 1988, p. 139). On this presentation, the amoralist is strikingly different from the normal moral agent, because he plays in a different league – he is outside the moral institution. Hence, as an 'outsider', he cannot make first-person genuine moral judgements and that explains why he is not motivated. Nonetheless, this option is not without difficulty, because if he is really an outsider why do the internalists care about his behaviour in the first place? Why does the internalist, at all, consider his rejection of morality as an absurd stance? This points to the fact that there is something that strong amoralism and morality share in common.[87] McNaughton rightly observes that:

> We cannot make sense of someone rejecting the moral viewpoint unless we can
> see him as embracing some other perspective from which morality is seen as of

---

[86] This also applies to the first group because in the spectrum of antisocial behaviors: some are extreme cases whereas others are mild.

[87] The internalist cannot dismiss the amoralist as a mere outsider, at least, the simple fact that he makes his decision in connection with morality shows something crucial about his stance. That is, his decision of rejecting morality as a whole or in part puts him in contact with moral values or systems. "In the context of the specific decision the individual believes that it is acceptable to place personal interest ahead of moral value" (Nisan, 1996, p. 76).

no value, or even as a positively undesirable institution…. We can begin to appreciate his objections, although we may think them exaggerated or even distorted, only because we have enough in common with his point of view to enter into it, even if we cannot share it. Two things follow. First, the simple characterization of the amoralist as someone who rejects moral values leaves amoralism totally mysterious. It is only when some suitable explanation is provided that we can begin to make this position intelligible. Second, it is very unlikely that the amoralist will utterly reject every moral value. In rejecting morality, he appeals to values that he shares with us. Not all of these need be moral values but it is likely that some will be. He may be opposed to a great part of morality, but it seems probable that what he seeks to replace it with will contain some elements in common with our present moral thought. (McNaughton 1988, p. 140)

Given this, the internalist charge begins to fade away, because the strong amoralist is possibly sharing similar cognitive state with normal moral agents, but he fails to be motivated because he rejects part of what morality requires. He might remain indifferent to moral concerns until the part of moral values that triggered his amoralist stance is replaced. Roughly, he suspends moral motivation. We might extend this explanation to the weak amoralist, he knows what morality requires, but sees no sufficient reason to be motivated by it. Possibly, he suspends his motivation awaiting the presence of sufficient reasons.[88]

Furthermore, let us stick with the case of weak amoralism. There is another reason why the internalist objection fails. Let us begin by making a distinction between *ontological objective experience* and *ontological subjective experience*. In the former, it is the view that we have epistemic capacity of knowing objective moral facts – facts that do not depend on our mental schemes. In the latter, we know the objective moral facts

---

[88] Possibly, it could be argued that such an agent is motivated by other value judgements he sincerely accepts. However, those other value judgements might be non-moral in content. Consider for example an agent who cares more about artwork. He might pay a huge sum of money for a piece of artwork regardless of the fact that there are many starving families living in his neighborhood. In other words, he might be morally indifferent if he does not see sufficient reasons to care about moral demands.

not in the absence of our personal capacities, traits, characters, history (such as moral upbringings and experiences). It would be odd to claim that we entirely mute these factors while engaging in ontological objective activities. As I understand it, moral realism is not claiming that in knowing moral facts we cancel out the first-person qualitative experience of the objective moral facts. Just as we cannot escape our shadows, we cannot avoid the first-person stance in our relation to the world. Again, it is not clear what the internalist means by the same cognitive state. Although we might have the same cognitive state by default, it is false to claim that qualitative states of the cognitive state are the same in all persons. Given this, the amoralist might share in the same cognitive state by default, but fail to be motivated due to his qualitative state. Notice that it is within the qualitative state that we recognize facts as objective reality. Further, suppose we idealize (that is, characterized it as a form of a single standardized state) the qualitative condition of the cognitive state needed for motivation, it would not only restrict moral motivation, but also nobody can fit into it, because it would be different from every other person's state. Alternatively, we cannot step into another person's unique qualitative experience.

Two things follow from this: First, it is one thing to know moral facts as objective reality in the world. This capacity is supposedly shared by normal moral agents in general; and this explain why virtuous or good persons, morally weak persons, morally wicked persons, bad persons or amoralists seem to accept, at least, the existence of moral demands. Second, it is another thing to know that we have unique first-person qualitative capacity based on our history, traits, situation, etc. It is possible to accept the truth of moral demands, without being qualitatively drawn or attracted to it because of the agent's (moral) upbringing, personal relation to morality, and the likes. The internalist seems to ignore this distinction. The amoralist can know that A is A, just as every other normal person, but A might appeal to him differently. What makes moral judgement genuine is because it represents things as they are in the world, not because it *appeal*s. On the one hand, to see, perceive or judge, and on the other hand, to be appealed or attracted to or to be motivated, are distinct features. The internalist seems to collapse them into a cognitive phenomenon, namely – moral judgement. In sum, these distinct factors come together to explain the link between moral judgement and motivation.

Chapter 7

A Defense of Externalist Moral Motivation

Virtue in rags is still virtue.

Hume, Treatise 3.3.1.19.

7.0 Introduction

Recall our definition of externalism in section 1.5, chapter 1:

> *Def1*: "[E]xternalism claims that the motivational force and rationality of moral considerations depend on factors external to the moral considerations themselves" (Brink 1989, p. 42).

Def1 is compatible with a range of metaethical theories. Suppose the claim of *Def1* is correct, we can plug in any theory that explains the role of thought and discourse in morality. For example, if the ideal observer theory is correct, namely, that the truth of moral judgements depends on the attitudes of a perfect hypothetical observer, then *Def1* holds insofar as the ideal observer theorist believes that moral judgement motivates agents contingently. The same applies to the various forms of cognitivism we discussed in chapter 3. Recently Kristján Kristjánsson (2013) argues that, contrary to the claim that Aristotle's virtue theory entails internalism, externalism is compatible with Aristotelian moral naturalism.[89] The truth of externalism does not hang on the plausibility of any form of moral realism or moral cognitivism. Further, Russ Shafer-Landau even argues that "externalism is true if noncognitivism is true, so long as there are any moral judgements that are expressive of non-cognitive attitudes that can fail to motivate" (Shafer-Landau

---

[89] Contrary to the claim which holds that, "Agents are practically irrational with regard to morality if their sincerely held knowledge of morality does not inform their action", he argues that "the moral judgement of continent persons does not motivate them intrinsically, yet the continent cannot be counted as practically irrational with regard to morality" (Kristjánsson 2013, p.440).

2000, p. 282). Hence, it would amount to a misconception to assume that the truth of externalism is determined by the plausibility of moral realism (Shafer-Landau 2000, p. 283). Given our discussion in chapter 3, I will defend a version of realist-based externalism.[90] We can realize our working definition by combining *Def1* with the cognitivist claim of moral realism:

> *Def2*: Externalism claims that moral judgements (that is, moral beliefs) do not motivate us in themselves; they only motivate us in connection with other external factors.

Def2 has the advantage of avoiding certain misconceptions that might arise from the scope of externalism. Just as in the case of internalism, externalism is defended on the fronts of reasons for action and motivation (Brink 1986, p. 28). Here we take care of this fact by narrowing the scope to a *motivational claim.* To be precise, our aim here is to propose a version of motivational externalism (hereafter simply externalism) that explains the widespread connection between moral judgement and motivation. I begin by presenting the externalist framework for explaining moral motivation (7.1). Since, my version of motivational externalism is based on the idea of moral agent as a moral person, I elaborate the distinctive ways in which morality is crucially significant to the human person, self and identity – moral identity (7.2). Further, I explore the two main construals of moral identity: Kantian-Kohlbergian and Blasian Perspectives (7.3). Drawing on Frankfurt's "Second-Order Volitions" and Blasi's "Moral-Self", I propose an externalist account of a widespread connection between moral judgement and motivation. On this view, motivation is based on moral identity. I argue that *moral care*, *moral-self* and *moral desires* are essential to moral motivation; and moral motivation does not stem directly

---

[90] I will not dogmatically rely on the plausibility of a given form of moral realism for the credibility of externalism: Our guiding principle is to conceive moral judgement as a cognitive state representing mind-independent truth-values of moral facts – facts that are natural but neither identical nor reducible to natural facts. This is our point of departure, of course, as Brink rightly points out; we are ready to abandon it should its metaphysics and epistemology be found wanting (Brink 1989, p. 24). Even in the case of readjustment, our cognitivist position would not threaten the credibility of our account of externalist motivation. However, this option is not available to realist internalism because its motivation claims are internally tied to its metaphysical and epistemological assumptions such that if the latter is false, the former suffers a knock-down attack.

from moral judgements; rather, it is from an independent desire, namely, a moral desire for the realization of the moral values we care about, hence, desire to be faithful to one's moral-self (7.4). Finally, I argue that this externalism is not condemned to explaining the motivation of virtuous persons in terms of desire *de dicto* as Smith alleges. In fact, it is possible, on the externalist construal, to account for a reliable connection between moral judgement and motivation in terms of *de re* desires (7.5).

7.1 Framework for Externalist Motivation

On the externalist framework, motivation follows contingently from moral judgements. It consists in factors external to an agent's moral judgements. This stance provides externalists with a wide range of possible factors for explaining moral motivation. On this view, motivation can be explained by appealing to *contingent psychological factors* such as conative states (for example, emotions – sympathy, compassion, empathy etc.), inclinations, self-sanctions, or desires in general (Roskies 2003, p. 52; Zangwill 2003, p. 134; 2008a, p. 91). One of the criteria for grounding motivation on these factors is that such factors are considered basic to human psychology. That is, they are "deeply seated and widely shared psychological trait[s]" (Brink 1986, p. 31; see also Nagel 1970, pp. 9 – 11).[91] For example, Hobbes' desire for self-preservation (1651, 1658) and Hume (1739) and Smith's (1759) sympathy are some of the examples of these psychological factors. To Hobbes' mind, the desire for self-preservation explains why humans are motivated to act in accordance with the demands of morality. "But still the motivational basis is prior to and independent of the ethical system, which derives from it" (Nagel 1970, p. 11).

Among other things, this implies that the source of motivation is independent of moral judgements; or that it is prior to moral agents' moral judgement, hence prior to moral systems. Appealing to such prior factors is meant to tie motivation to factors that are basic to human conditions. In Hobbes' case, the presence of desire for self-

---

[91] There are various ways of defending motivational externalism, in most recent accounts; motivation is construed as contingent on psychological factors that can be known *a posteriori* (Brink 1989, p. 28).

preservation is primitive to humans, hence, prior to moral systems. However, it is good to note that the externalist does not claim that these factors *always* provide agents with motives for acting morally (Svavarsdóttir 1999, p. 161).[92] It allows room for the possibility of motivational failures, for example, in cases where other values or self-interests outweigh moral considerations.

Notice that the externalist appeal to independent desires or any other non-cognitive factors is not only compatible with Humean psychology of motivation, but with anti-realist internalism as well. Recall Hume's constraint on reason:

> Morals excite passions and produce or prevent actions. Reason of itself is utterly impotent in this particular. The rules of morality therefore are not the conclusions of our reason. (THN 457/294)

In his view, passions are the source of motivation; and given the inertia of reason (and its product – beliefs), moral judgements motivate only in tandem with pre-existing corresponding desires.[93] In addition, anti-realists appeal to factors based on conative states (for example, emotions, feelings, self-interests, inclinations, etc.) to explain the widespread motivation in moral agents. However, the anti-realist parts way with the realist externalist by collapsing the act of judging and the conative state into a complex mental state, such that, to judge, on this view, is to express a conative state. For example, to judge is to express a sympathetic response; and if an agent accepts the content of her moral judgement, she is necessarily motivated. Apart from the claim that the anti-realist internalist does not take motivational failure seriously,[94] by rolling moral judgement and

---

[92] Hobbes' intuitive appeal on why we are moral is based on the claim that the desire for self-preservation is basic, and as such, morality serves only as a system that promotes this end.

[93] It would not be false to say that most of the externalists are moral realists and Humeans about motivation (Brink 1997; Svavarsdóttir 1999, 2005; Railton 1986; Zangwill 2003; Strandberg 2007).

[94] Shafer-Landau rightly observes that it is a misconception to think that anti-realist internalism tightly secures the necessary connection between moral judgement and motivation, because some of the conative factors, which agents express in their moral judgements, might not be intrinsically motivational. "The lesson here is that the connection between noncognitivism and internalism needn't be as tight as many have thought. The connection is insecure if we fail to identify the candidate states that are said to receive expression in moral judgements. The connection may sever even if a candidate is identified, since it may be the case that some non-cognitive states fail to be intrinsically motivating. In the end, internalism will receive support from noncognitivism only if it successfully discharges both of those argumentative burdens" (Shafer-Landau 2000, pp. 280 – 281).

conative state into a lump, he seems to eliminate the role of thought and discourse in morality. "Philosophers who believe that there is no room for rational assessment of the basic springs of motivation will tend to be internalists, but at the cost of abandoning claims to moral objectivity" (Nagel 1970, pp. 7 – 8). Hence, realist motivational externalism gives us the advantage of avoiding a black and white explanatory model of psychology of moral motivation. Unlike realist internalism that relies essentially on cognitive state, and anti-realist internalism that relies solely on conative states, realist externalism appeals to both states, thereby capturing a bigger picture of psychology of moral motivation. This posture enables externalism to explain the ubiquitous motivation as well as motivational failures in moral agents.

In addition to these psychological factors (external to moral judgements), moral realism holds that there are *features that make actions right*. This idea is closely tied to the notion of correctness of actions we saw in chapter 1: There are features that make some actions better than their alternatives. Suppose we have two sets of actions, A, with g, s, k features; and B, with x, y, z features. First, we can distinguish these actions based on their features; and second, based on their respective features, we can say that one of the sets of actions is preferable to the other. In other words, motivational externalism requires that moral agents hold the view that some actions have features that make them right or preferable in comparison to their alternatives. Notice that the content of these features is neutral at least in two ways. First, it does not tell us whether the features are moral or non-moral features; and second, it is silent about what determines the rightness of those features. To the former, it might be asked: Is it odd to claim that some natural features essentially figure in moral actions? I don't think so, because some of them such as caring for one's family, keeping promises, concern for people's welfare, can make actions right (Strandberg 2007, p. 8). However, this is not to place non-moral features on par with moral features. Essentially, it is thought that what makes moral actions right are moral features. That is, moral actions have features of being morally better than their alternatives (Zangwill 2003, p. 144).

Nevertheless, in the metaethical domain, what determines the rightness is perceived differently. For example, recall the moral constructivist position, that moral

161

correctness is the product of rational procedural construction. Thus, on this view, what makes an action right consists in features that result from rational constructions. Suffice it to say that the answer to 'what make actions right' differs across the various forms of moral cognitivism. I will not pursue these various construals further; let us focus on non-reductive naturalist notion of rightness. It is in order to recall that this form of moral realism holds that although moral facts supervene on natural facts, they are not reducible or identical to them. On this construal then, the features in question are moral features supervening on natural features.

Furthermore, the third element that externalists need for explaining motivation is *belief.* Moral agents have to believe that actions in question have relevant features. Suppose it is right to perform A-action. While driving down the road an agent 'C' encounters an incident and he wants to do 'A'. On this view, he has to believe that 'A' has the features that make actions right. By acting accordingly, he will be satisfying the features of A-action. It must not be forgotten that agents can be mistaken in their beliefs. They might fail to hit their targets, that is, to capture the relevant features that make actions morally right. For example, C might realize that his action does not capture the right-making features of A. He might change his outlook and believe that an alternative action can better satisfy the features of A. However, it should be noted that change of belief is not necessarily due to mistake. It might as well be that the original desire to perform the initial action clashes with other interests or values. Notice that in order to discover moral rightness, that is, 'what makes action right', agents seem to possess advanced or detailed moral theories (Kristjánsson 2013, 424), however, not in all cases, because, as we noted above, natural features might figure in moral actions (Strandberg 2007). With these elements, externalists can explain moral motivation as follows.

Motivational Externalism (ME): An agent who believes that an action φ is right is motivated to do φ because he holds the view that φ has features of being morally right and he has the relevant independent desire to perform φ.

The core claim here is that *moral motives* for acting on moral judgements are based on independent and distinct desires; and the *moral desire of being true to oneself* is one of them. Shortly I will offer an externalist explanatory account that grounds moral

162

motivation on moral desire driven by our care for morality. We are morally motivated because we care about morality – *Moral Care*. In essence, I care about moral value(s) and desire their realization, because 'who I am ' matters to me. Before that, let us consider the notion of personhood and the role of morality in personal identity.

7.2.   Personal Identity as Moral Identity

"What (or who) is a person?" is without doubt one of the widely discussed philosophical questions.[95] Scholars (such as Strawson 1959; Frankfurt 1971; Taylor 1989; Watson 1987; Raz 2006; Velleman 2002, 2005; Westlund 2011; Schechtman 1996, 2001, 2007; Goldie 2003, 2012; Baker 2000; Meyers 2012, 2014) have offered different answers to the question about the nature and constitutive properties of personhood. For the purposes of this work, I will not directly delve into these various views of personhood. I will, rather, focus on Harry Frankfurt's account for two main reasons: First, his construal of personhood and agency neatly captures the agentic capacity for forming higher-order desires, which includes the ability of reflecting on existing desires; and whether an agent wants the desires to be actually action-guiding. Externalism has better resources for analysing this capacity of personhood and agency, precisely, how it figures in the connection between moral judgement and motivation. Second, given that the efficacy of motivating reason depends on how morality fares while an agent engages in this evaluative structure, Frankfurt's construal offers us a framework for explaining the complexity and different possibilities, such as amoralism, which are involved in the evaluative structure of agency in individual persons. Accordingly, in this section, I draw the link between *who is a Person?* (Personhood) and *who am I* (Personal Identity)? My objective is to motivate the claim that morality largely shapes 'who we are', hence, *Moral Self Identity*.[96]

---

[95] There are also legal as well as psychological works on personhood (e.g. Martin et al. 2010; McDonald 2007; Ikäheimo and Laitinen 2007; Martin and Bickhard 2014; Jungert 2014; Rose 2010; Pribytkova 2009; Lindroos-Hovinheimo 2015; Travis 2015; de Mori 2001; Boyle 1979; Groarke 2010).

[96] I will use personhood and selfhood (as well as person and self) interchangeably; however, this is not to deny the difference that might exist between them.

Frankfurt's concept of personhood has its mainstay in desires. The starting point of *Freedom of the Will and the Concept of a Person* (1971) was Strawson's ascription of consciousness and corporeal properties as essential to the concept of a person. On Frankfurt's view, such an ascription is a misappropriation, hence, problematic because there are other entities besides humans who possess both conscious/mental and corporal/physical properties. He argues that:

> Our concept of ourselves as persons is not to be understood, therefore, as a concept of attributes that are necessarily species-specific. It is conceptually possible that members of novel or even of familiar nonhuman species should be person; and it is also conceptually possible that some members of the human species are not persons. (Frankfurt 1971, p. 6)

We can draw at least two main implications from Frankfurt's claim. First, the concept of personhood is not reducible to physical and biological life (see also, Taylor 1989; Sullivan 2003). This position rejects naturalistic construal of personhood. Naturalism about personhood rejects phenomena (such as feelings, attitudes, values, desires) within the scope of subjective experience, because they are considered indescribable in 'absolute terms' (Sugarman 2005). However, accepting the naturalistic view amounts to significantly downplaying, if not terminating, our rich subjective-inner life or what Taylor called 'subject-related phenomena', which include human meanings, values, desires, needs, etc. Second, the concept of personhood is not merely a tool for separating nonhuman species from humans. Rather, it serves to explain the attributes which are essential to our most human concerns. Alternatively, if an agent does not possess the essential attributes in question, irrespective of being a member of the human species, she does not qualify to be called a *person*.

Thus, Frankfurt argues that desires are at the heart of personhood. However, he quickly notes that desires *per se* (including motives and choices) are not definitive given that nonhuman creatures possess such capabilities, even if to a certain extent. Hence, the *will* is required. In other words, a person is an agent who engages her will in her desires. The desires in question are what Frankfurt calls the second-order desires, as opposed to the first-order desires.

Many animals appear to have the capacity for what I shall call "first-order desires" or "desires of the first order," which are simply desires to do or not to do one thing or another. No animal other than man, however, appear to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires. (Frankfurt 1971, p. 7)

From the foregoing, it is not just the members of nonhuman species, but humans as well, who are capable of having first-order desires. First-order desires, for example, 'I want a cup of water', do not admit additional desires. Besides, the possessors of first-order desires seem not to be in control of their desires. That is, they appear to be responding to some sort of natural stimuli. This idea is akin to Kant's account of an agent acting within the scope of phenomenal self. I will return to this idea shortly in section 7.3.1.

Nevertheless, in the second-order desires agents desire their desires: Their desires have desires as objects. According to Frankfurt, second-order desires come in two main flavours: "[E]ither when he wants simply to have a certain desire or when he wants to have a certain desire to be in his will" (Frankfurt 1971, p.10). In the former, the agent *wants* to *desire*, say, to stop smoking cigarette, but this desire hinges on a first-order desire, namely to stay healthy. However, this form of second-order desire might fail to move the agent all the way to action if it is in conflict with or dislocated from the first-order desires. In the latter, namely "second-order desires volitions or volitions of second-order desires" the agent cares about his will. Here the agent possesses a self-evaluative power of thinking and reflecting upon her "desires and motives, to form judgements and desires with respect to them" (Lapsley 2007, p.4). She wants certain desires to be part of her will, that is, she wills the desires in question "all the way to action" (Frankfurt 1971, p.8). Thus, Frankfurt calls the possessor of second-order volitions a *Person*, as opposed to a *Wanton*. A wanton: "Not only does he pursue whatever course of action he is most strongly inclined to pursue, but he does not care which of his inclinations is the strongest" (Frankfurt 1971, p. 11). Hence, we can classify agents with either first-order desire or second-order desires without volitions as nonpersons or wantons, because they do not care about their will or the desirability of their desires.

Following Frankfurt, Charles Taylor (1989) argues that a person is defined by the level of her engagement with the things or values that are significant to her. She *strongly evaluates* those things, commitments, values, desires, etc., within a 'horizon of significance'. An agent's level of engagement is determined by her connectedness to the values or goodness she sees in things etc. (Meyers 2012, p. 157). Assuming Taylor's claim is plausible, it can be argued that the agent's connectedness is driven by her desire to *know who she is* – Personal Identity. He writes,

> My identity is defined by the commitments and identifications which provide the frame or horizon which I can try to determine from case to case what is good or valuable, or what ought to be done or what I endorse or oppose.(Taylor 1989, p.11)

From the foregoing, we can impute to Taylor the following thought: The values, goods or ought to be doneness that a person endorses is an essential part of the Self, such that he aspires and desires to conform to them. Conversely, a person shuns the things he disapproves, because they are not in consonance with the self. In other words, when I desire my "desires to be a part of my will" or when I connect myself to "strong evaluated values, commitments and desires within a horizon of significance", I do so because my personal identity matters to me. I care about them because I care about 'who I am' (or possibly, the kind of person I want to become).

More so, the notion of personal identity, at least, straddles metaphysical and moral discussions. Although, for our purposes we shall focus on how personal identity relates to morality; however, let us very briefly explain the metaphysical aspect. For example, it is often asked: What centrally defines who we are? John Locke champions the idea of memory, precisely autobiographical memories, as crucial to personal identity. On this view, it is memory that makes us the *same* as the *persons* we are in the past (see also Perry 2002). Unlike Locke, David Velleman essentially construes personal identity as self-conception, where an agent represents herself as herself (Velleman 2002). Sugarman argues for a concept of personal identity embedded on a historical *cum* sociocultural constitutiveness (Sugarman 2005). However, although memories, mental faculties, embodiment, and sociocultural factors, play roles in our personal identity (Sugarman

2005; Velleman 2002; Blok, Newman and Rips 2005; Nichols and Bruno 2010; Parfit 1971; Shoemaker 1959), it is argued that dispositional traits such as desires, preferences and personality (Gelman, Heyman and Legare 2007; Haslam et al. 2004; Allport 1937) are crucial to personal identity.

Furthermore, there is an ongoing strong contention that morality heavily shapes our personal identity and self-perception (Stohminger and Nichols 2013; Parfit 1984; Skitka, Bauman and Sargis 2005; Goodwin, Pizza and Rozin 2014) to the extent that morality is more crucial to identity than memory. Stohminger and Nichols strongly argue that "the self is not much the sum of cognitive faculties as it is an expression of moral sensibility; remove its foothold on that world, and watch the person disappear with it" (Stohminger and Nichols 2013, p. 169). Taylor argues that "being a self is inseparable from existing in a space of moral issues" (Taylor 1989, p. 112); and Carr believes that "although there are other senses in which human agents may be regarded as persons, the most significant sense in which they are persons is that in which they are moral agents" (Carr 2001, p. 82). Nevertheless, while it would be too ambitious to dismiss the roles of the mind or other aspects mentioned above in a few strokes.[97] Given that we usually consider each other at least as moral agents, we are allowed to say that morality, indeed, plays important roles in our social practices and self-understanding. However, we need to further explicate the role of morality. First, morality plays a role based on the capacity of personhood. Generally, we consider a person, and not a wanton, as a moral agent. An agent is conceived as a moral agent just in case she is a person. Notice that personhood underlies moral agency. Alternatively put, once an agent meets the requirements of personhood (namely, being considered as a person), the capacity of taking up a moral perspective comes with it. This explains why we normally consider each other not just as persons, but also as moral agents. Second, morality also plays a role based on personal identity or personality. On this level, a person (understood as a moral agent) faces a practical question, namely whether she wants to act morally and whether she wants to be a person for whom morality plays a crucial role in her self-understanding. Notice that

---

[97] I think that it is wrong to claim that the idea of moral self is possible in the absence of mental faculties. These faculties help in providing the perceptual and conscious apparatus required for moral perception or judgements; and they are relevant in understanding of the concept of self or personhood.

answering this question in terms of wanting to be a moral person or identifying with morality's demands is merely a possibility. Hence, while persons remain moral agents in general, i.e. they remain *qua* persons capable of moral agency, this does not imply that they also adhere to morality's demands or identify with them in terms of shaping their personality. Assuming the possibility of the second role of morality; and that of transitioning from the first role to the second role of morality hold, we can argue that morality can form an important block with which we construct our identity. It can give our personhood an identity, namely, a moral identity. Moral identity lies at the very core of our being, and it plays a crucial role in our self-definition and self-recognition. Moral identity is construed differently. However, we shall briefly consider it here from two main perspectives: Kantian-Kohlbergian perspective and Blasian perspective.

7.3. Moral Identity
The Kantian-Kohlbergian Perspective

It goes without saying that the notion of self has a place in Kant's theory of moral agency as well as moral rationality (Carr 2001). Kant divides the self into phenomenal self and noumenal self. Roughly, the former is constrained by space, time, and natural requirements, whereas the latter is not constrained by them. As organisms, natural elements such as inclinations, desires, passions, impulses, are significant determinants of human activities. The phenomenal self, on this view, is governed by these natural elements. To Kant's mind, these elements would be problematic to moral agency, because they constrain, thereby making moral actions not free (Johnson 1993). Hence, he understands moral-self as a noumenal self – the self that is not constrained by the above-mentioned natural elements. As Lapsley puts it: "The noumenal moral agent can will purely, in complete freedom of the contaminating influence of passion and the determinisms of sensible experience" (Lapsley 2007, p.2). On this view, Kant seems to conceive a moral agent as 'sterilized-self', the sort that is detached from "deep human nature" and personal particularities – personal identity, emotions, dispositions, traits, inclinations (Zangwill 2003; MacIntyre 1984).

Similarly, Lawrence Kohlberg (1969) does not place much emphasis on the role of 'embodied' personhood in moral agency. This is partly due to his Kantian commitment;[98] and partly due to the structure of moral developmental stages he inherited from Jean Piaget. To Kohlberg's mind, the self enters the domain of moral action only when the moral agent judges his action as satisfaction of moral sanctions. However, the self is understood here as sign of attaining a higher cognitive development. In other words, to be a Kohlbergian moral agent is to attain the stage of full comprehension of moral prescriptions and principles.

> At the highest stages one has a better appreciation that moral principles make prescriptive claims upon the self; that moral principles oblige the self to enact what duty requires. For Kohlberg, then, it is clear grasp of prescriptivity that launches the responsible self into action. (Lapsley 2007, p.3)

On this basis, what constitutes moral agency is the extent to which one grasps moral principles and commits them in the line of his action. In sum, it is the understanding of the sanctions of moral principles, and not the self with its dispositions, traits etc.[99] that are necessary for moral agency.

Blasian Perspective

The focus of Augusto Blasi's work was the gap between selfhood, identity and moral agency in moral psychological literature. Precisely, he seeks to overcome the limitations of Kohlbergian cognitive moral development by bringing 'deep human nature' – the self, personal identity, desires, dispositional and character traits – and morality together. On this view, it is no longer the sanctions of moral principles embedded on fully developed cognitive faculties, but self-sanction, rooted in the moral person, that drives moral agency. Conversely, an agent might grasp the demands of moral judgements, but if moral

---

[98] Kantianism is based on the thought that the will must be freed from the sensory experiences, bodily desires such as inclinations, passions, emotions etc. These forces are considered to be antagonistic to reason, such that we are better-off keeping them outside the frame of moral agency (Johnson 1993).
[99] Kohlbergian attitude treats the role of self or personhood in morality as suspect. Lapsley puts it thus, "deeply personological constructs were viewed as obstacles to mature moral deliberation, or as sources of bias and backsliding that had to be surmounted by the rational moral agent" (Lapsley 2009, p. 240).

ideals or values are not central to the agent's self (that is, if they are not at the heart of his personal identity), he would see those moral demands as foreign bodies. In other words, it is documented that the "moral judgement-action gap" results from locating moral agency heavily on cognitive faculties – deliberative moral reasoning and Socratic-Kohlbergian 'knowing is doing' principle (Cervone and Tripathi, 2009; Frimer and Walker 2008, 2009; Walker 2004; Lapsley and Hill, 2009; McAdams, 2009).

In his concept of 'Self Model', Blasi (1983, 1984) essentially aims at closing this gap by placing moral identity and personhood at the core of moral agency. Thus, not only to explain widespread moral motivation, but also differences found in agents' moral behaviours.[100] According to Blasi, moral identity is the extent to which morality, precisely, moral ideals, values, traits, desires, are central to agent's personality or character (Blasi 2004; see also Colby and Damon 1992). The difference in moral behaviours is largely due to the various levels of centrality of moral values, traits and ideals in an agents' self-definition (Blasi 1995). It is the level of the rootedness of morality to one's self that determines one's moral commitments and motivation. In some people, the level is high, whereas in others, it is moderate or low. Also, in some people, there are certain moral values and moral desires (such as being good, compassionate) that occupy central place in their self-conception, whereas in others, it might be fairness, justice (Blasi 1984). More so, Blasi argues that human beings have the desire for consistency; and self-consistency is essential to agents' being (See also Erikson 1964). "This tendency provides the motivational impetus for moral action, so that a person whose self-definition is centred on moral concerns will feel compelled to act in a manner that is consistent with his or her moral self-construal" (Shao, Aquino and Freeman 2008, pp. 514 – 515; see also Blasi 1984; Colby and Damon 1992; Oliner and Oliner 1988).

---

[100] It is also good to note that moral identity is studied from the social-cognitive perspective, the view that "a person's moral identity is stored in memory as a complex knowledge structure consisting of moral values, goals, traits, and behavioral scripts" (Aquino et al. 2009, p. 124, for more discussion see Aquino and Reed, 2002; Lapsley and Narvaez, 2004).

7.4. Moral Identity-Based Motivation
Moral-Self

'Who I am' and 'What I do morally' are closely tied; and this is seen in how we *relate* ourselves to morality (and how others attribute such a relation to us). This relation is crucial for understanding personhood in terms of the capacity for developing a personality as well as shaping our self-understanding. Morality colours our persons, such that we can be identified based on our 'moral signature'. Moral signatures are those moral traits, ideals, values that are central to the moral person. Given that we set our preferences differently, our level of moral commitments varies: A fact which essentially hinges on the centrality of morality to one's self. Suppose, for example, that 'promise keeping' is an ideal very central to Ann, a tennis player; and there is an important game going on. But she has already promised her grandma driving her to the dentist for a check-up. Given that keeping her promise is at the core of her self-understanding, she is motivated to act accordingly – fidelity to one's moral identity. "One has a moral identity when one strives to keep faith with identity-defining moral commitments, and when moral claims stake out the very terms of reference for the sort of person one claims to be" (Narvaez and Lapsley 2013, p. 243).

Being morally faithful is a form of reaffirmation of personal identity. Self-fidelity is a personality mechanism that engraves moral ideals and values deep to 'who we are'. It makes our moral signature authentic and visible, such that 'who we are' is consistent with 'what we stand for morally' and 'where we stand morally'. Conversely, to the moral agents who identify with morality based on Frankfurt's second-order volitions, namely, morality as essential to the self, cases of moral self-infidelity will dislocate them from the moral-self. For example, by failing to identify with her core values, Ann separates herself from herself. On this view, being unmotivated by moral demands is not a moral principle-betrayal, but a moral self-betrayal. Again, if we take morality to be essential to a person's self-understanding, that is, to who we are and want to be, then moral self-infidelity threatens to trigger off a loss of personal identity. Suppose an agent responds to moral demands, but her action does not spring from her moral self, although, such actions might

171

be thought to follow from moral demands, however, they do not sufficiently meet the requirements in terms of actually being motivated to act morally. Simply put, mere conformity to moral demands is different from being motivated by moral reasons. In the case of the former, the agent's actions are divorced from her person. However, the divorce I have in mind is the following: mere moral conformity separates the agent from her overall moral self-understanding. This distinction is in order, because an agent might act from reasons stemming from egoistic character trait. In that case, the agent's action is not divorced from her person completely, but only from the moral part of her self-understanding, which she takes to be crucial for her self-understanding. However, in the latter case, moral actions stem from an agent-as-whole-person. In other words, moral functioning is not to be restricted to moral demand-based judgement at the expense of moral agency of a whole person (Walker 1999). In the Moral-Self model, the moral motivation is holistic, that is, the desire for being moral stems not from external forces of moral demands, but from the self's commitments to morality.

Care and Moral Motivation

The place of morality at the core of one's self determines the intensity of his moral commitment. However, irrespective of our moral commitments, if we do not *care* about morality, we will remain unmotivated by these values, principles, standards etc. The capacity for caring is essential to moral motivation. This does not only apply to morality but also to non-moral cases. For "what moves us is that there is something of which it is both more precise and more pertinent to say that we c*are about it* or that we *regard it as important to ourselves*" (Frankfurt 1999, p. 155). The emotional characters of caring differ from moods. The latter has a lesser affective intensity, and unlike emotions, they are not triggered by objects, events or situations (see Frijda 2009; Weiss and Cropanzano 1996; Pathak et al. 2011). Emotions are episodic and intensive-specific affective phenomena, such as anger, fear, sadness, happiness etc., caused and directed to someone or something – object, event. Unlike, basic emotions, Agnieszka Jaworska rightly argues that caring is a form of complex emotion.

> [C]aring has an even more complex structure than most ordinary emotions—it is best understood as a structured compound of various less complex emotions, emotional predispositions, and also desires, unfolding over time in response to relevant circumstances. (Jaworska 2007, p. 483)

Emotions, both basic and complex, are characterised by the idea of *directedness*. Emotions are not only triggered by objects, situations etc., they are directed to them as well. For example, in the case of biblical Cain and Abel, first, seeing that God favoured his brother's sacrifice, Cain became envious (situation or circumstance that triggered his envy). Second, he was envious of Abel (it was directed to him). In case of caring, a fan is happy when the football team he cares about wins (the 'win' triggers, whereas, his care is directed to an object, namely his team). The directedness of care towards an object or someone, roughly, is captured etymologically in 'moving-outness' of emotion (*emovere*). When an agent cares, she expresses an emotional attitude that *moves toward* objects, values, preferences, interests, needs, etc., out there, which are *important* to her.[101]

Furthermore, the directedness of care points to another important element, namely its intentionality. Given that it is triggered by specific objects, situations or contexts, the act of caring exhibits the *aboutness* of something or someone. Jaworska argues that "complex emotion is, as a result, *about* the object" (Jaworska 2007, p. 484). For example, our neighbour's child cares *about* her Lego toy in a way that differs from how he cares *about* pudding. The moving-outness of care is specifically directed, because it captures the objects or situations in question neatly. In other words, the character of specific-directedness of caring drives our *concern* for the objects or situations in question, thereby explaining why we care differently about different things, values, interests etc. However, this is not to deny cases where our care about something or someone clashes with other sets of concerns. In such cases, the directedness as well as the aboutness of objects is blurred or confused.

---

[101] Frankfurt rightly observes that a mere want of preferring, say, skating to skiing does not mean we care about those we prefer to others. "It is quite common for people to want various things without actually caring about them, and to prefer satisfying one of their desires rather than another without regarding the object of either desire as being of any importance to them" (Frankfurt 1999, p. 157).

Additionally, morality, for some people, belongs to the class of values, principles, standards that we care about. Their importance is paramount that even instances of ignorance do not affect their ontology. The fact that some objects or values are less known (either at a given time or place) does not make them less important or valuable in themselves.[102] In other words, that morality, given its normative characters, is value-factually[103] different from, for example, physical objects, does not make it less important. On the contrary, morality occupies a pride of place in the litany of things or values we care about. Frankfurt eloquently captures this issue as follows:

> Surely there are certain things that are inherently and objectively important and worth caring about, and other things that are not. Regardless of what our own desires or attitudes or other mental states may happen to be, surely there are some things that we should care about, and others that we certainly should not care about. Is it not unmistakably apparent that people should at least care about adhering to the requirements of morality, by which all of us are inescapably bound no matter what our individual inclinations or preferences may be? (Frankfurt 2006, pp. 20 – 21)

Essentially, we care about morality because we care about the kind of persons we are or want to be. Michael Kühler rightly argues that "[normative] claims not only tell us what to do, but, by that, also what kind of person to be. Hence, they comprise at least a partial answer to the question of who one is or wants to be and cannot be ignored or failed to be met without running the risk of compromising one's self" (Kühler 2012, p. 191). Given the possibility of the second role of morality, that is, identifying with morality in terms of Frankfurt's second-order volitions, we can say that our personal identity is closely tied to moral identity, because the latter immensely shapes the former. When moral values are at the core of our self, we strive to be faithful to our moral-selves. Nevertheless, no matter the place of morality in the self, if we lack the capacity for caring about morality (what I refer to as moral care), our moral motivation would be blocked. The capacity of moral

---

[102] What people do not care about may nonetheless be very important to them, obviously, because of its value as a means to something that they do in fact care about (Frankfurt 1999, 1998, 2006).
[103] We already considered this issue in chapter 2.

care is a crucial raw material in the economy of moral motivation. In other words, while the centrality of morality in the 'self' determines the intensity of moral commitments, our care about morality underlies our motivation to act morally. Our friend Olivia is a case in point. Although she (merely) acknowledge moral demands as justified, she remains unmoved, because she does not care about them. Precisely, she no longer cares, identifies or feels committed to moral demands and moral reason in a way that is related to her self-understanding.

Of course, we have to note that moral care (in fact cares in general) comes in various degrees. It is possible for someone to care about morality in an unqualified manner, but it is, at least, conceptually possible that the degree of his moral care is stronger in certain values than in others. For example, keeping promises might be at the core of Ann's moral-self and she cares about it. However, it is possible that she cares less about justice. Hence, suppose she encounters a case of injustice. She might be unmoved or indifferent given that she cares less about justice. In addition, there are cases where agents care less about morality in general. In such cases, it might be that morality is not central to them or that morality is central to them but their capacity for care is impaired or deficient (for example, as in the cases of psychopaths and acquired psychopaths).[104] In addition, it is possible in cases when moral concern does not matter to them (for example, Zangwill's mercenary in chapter three) or when other non-moral interest override moral concern, as in the case of a wealthy king who cares about the beauty of his palace over his starving subjects. In other words, we care differently about morality (also about the various moral principles and values). Thus, the degree of our moral care determines the strength of our moral motivation.

---

[104] Consider Damasio's (1994) patient Elliot who damaged his ventromedial prefrontal cortex (VMPC). It is documented that he was, after his injury, no longer himself, his personal relationships, plans were immensely affected. "He was incapable of sustained goal-directed behavior in any aspect of his life. And yet he was remarkably emotionally unaffected by his tragic life history …. Elliot did not seem to care about himself. He did not experience joy or satisfaction when he was flourishing or frustration over his misfortunes; no anger at agents who caused his misfortunes; no pride in his successes or disappointments over his failures. He had no sustained desire to help ensure those successes or to help avoid the failures" (Jaworska 2007, p. 486). It could be argued that given that his impaired emotional responses and feelings, he is not only unable to see himself as himself, identify and commit himself to certain values, but also he can no longer care about himself. In other words, he cannot care about his moral-self (suppose he still has one).

More so, in caring there is a persistent desire to see the realization of those things, values etc. that are important to us. This shows that some sort of dispositions is required, that is, a stable disposition to sustain the directedness. In the case of moral care, the stable disposition sustains the directedness to morality, which in turn secures "a pattern of emotional investment retained over time" (Jaworska 2007, p. 487). We cannot (morally) care if our dispositions of a sustained directedness (towards moral values or principles) over time are unstable. Further, in the moral context, the persistent desire is, using Frankfurt's terms, a moral desire to desire the values at the core of our moral-self all the way to action. This is not to claim that moral desires are not manifestable in forms of first-order desires or second-order desires, but as we saw above such first-order desires are not the type we ascribe to persons, but to wanton (I will return to this shortly in our discussion about *de re* and *de dicto* desires).

Moral desire, as I understand it here, though distinct, is closely tied to moral care, such that when I care about morality, I (morally) desire it to be realized. Consider the case of Kirk Noble Bloodsworth who was convicted of rape and first-degree murder of a nine-year-old girl. He maintained his innocence from day one, irrespective of eyewitnesses' testimonies against him. Assuming that justice is at the core of his mother's moral-self: She cares about it and desires it to be realized in her son's case given that she knew that he was innocent of the charges. Imagine both the frustration or sadness she felt when Kirk was incarcerated; and the joy or satisfaction she felt when he was freed after nine years. In other words, when we (morally) care, we desire it; and its realization causes feelings of joy or satisfaction in us, otherwise, frustration or disappointment (Jaworska 2007; Frankfurt 1999). However, by caring about promise keeping, Ann desires it to be realized, assuming she fails, she not only feels frustrated, but also self-betrayed (which could lead to self-alienation or loss of moral-self in a case of chronic frustration due to a repeated moral self-infidelity).

In other words, moral care and moral desires, tightly work in tandem, determine both the desirability and the agents' yearnings for the realization of the moral values in question over other non-moral considerations. It is not merely the moral values and ideals at the core of one's self *per se* that motivate. However, we are not essentially arguing that

176

first-order desires do not motivate us to act accordingly, regardless of whether we (also) care about the object (or the desire) or not. What we argue is that *internalism* seems to *exclude* the possibility of caring about morality. If we, for instance, acknowledge that at least some people care about morality (that is, taking it as a very important part of their self-understanding), we are allowed to say that the moral desire to be faithful to those moral values and ideals is their primary motivation to act accordingly. For example, promise keeping might be central to Ann's self-conception, but if she cares less, that is, if the (moral) desire to keep faith to her promise is weak, the drive to be in unity with the self reduces. This in turn affects her motivation accordingly. Moral care and desire unite the self and moral ideals and values into a whole, hence, the Moral-Self. They not only heighten an agent's moral commitments, but also, they make the moral person to yearn and act in a manner that is *ad rem* to his moral identity. Using Frankfurt's terms, a moral person cares about the desirability of desire to be good or moral; and she *wills* it so that her actions will be in unity with her moral-self (see also Blasi 2005; Nisan 1996).

7.5. Motivational Externalism and Widespread Motivation

On the Moral Identity-Based Motivation, moral care, moral-self and moral desires are essential to moral motivation. Motivation does not stem directly from moral judgements. Rather, it is from an independent desire, namely, a moral desire for the realization of the moral values we care about, hence, the desire to be faithful to one's moral-self. At this point, let us recall Motivational Externalism (ME*):

> ME*: An agent who believes that an action ϕ is right is motivated to do ϕ because he holds the view that ϕ has features of being morally right and he has the relevant independent desire to perform ϕ.

Suppose that Peter judges that to save drowning immigrants is right, he is motivated to act accordingly because of his strong sense of moral-self and the belief that the action in question has features of being morally right. Moreover, Peter has a moral desire (embedded on moral care) to save drowning immigrants. On this view, it is the centrality of morality in the agent's personhood and intensity of moral desires, working in close

177

tandem with moral care, that track moral motivation. If, for example, Peter places saving lives above other considerations, and he cares very much about the moral issue in question; and it is this distinct moral care as well as the moral desire to see its realization, that gives him the motive to act on the moral judgement: 'Saving drowning immigrants is right'. The interplay of moral self, moral care and moral desires explain not only motivation, but also the differences in moral motivation (Blasi 2004a; see also Hardy and Carlo 2005).

It is at this juncture that *moral education* and *self-creation* play crucial roles in moral agency (Shafer-Landau 2000). The level of agents' moral upbringing contributes to the formation of moral identity as well as imbues on them relevant desire for moral concerns. So also agents can acquire good sense of moral-self and desires by essentially determining and placing morality at the core of their personality or character. In the same vein, we invest in the things or people we care about. We want to protect and keep them even to the extent of sacrificing heavily or losing our lives, because they are important to us. Consider, for example, a loving widow who cares about her four children and promotes their well-being at the cost of her comfort. Since morality is a crucial aspect of who we are, it requires some sort of investment. The sort of investment I have in mind is captured in the words "nurture" or "taking care of". Given that 'who we are' and 'what we do morally' are in a constant dialogue, we are required not just to cultivate our moral care, but to foster its growth as well. Roughly, 'nurtured moral care' serves as cement that binds our personal identity to our moral-self.

That said, agents who are faithful to their good moral upbringing are motivated accordingly, given the centrality of morality to their moral-self, 'nurtured moral care' as well as the intensity of the desire to keep faith with their moral identity. However, agents with not-so-good moral upbringing or deficient moral education seem to lack these elements; and this explains the cases of motivational manifestations as well as motivational failures in moral agents. In other words, in acting consistently with one's moral identity, an agent is not only being true to his moral self, but he is also making his moral signature visible to other agents. He expresses those signatures or features that

make him stand out as a unique moral person. On this basis, the externalist can provide a widespread phenomenon of moral motivation.

7.6.   Moral Fetishization: When Moral Desire becomes Bad

Recall Smith's charge against externalism, namely, that externalism cannot explain the reliable connection between moral judgement and motivation. According to Smith, a change in motivation directly and reliably follows from a change in moral judgement in good and strong-willed agents. Smith strongly claims that such a reliable explanation is not available to motivational externalism. To his mind, the most externalism can offer is a motivation stemming from *de dicto* desires. According to him, good and strong-willed persons, at least, are motivated by *de re* desires. Smith accuses the externalists of fetishizing virtues, thereby making the motivation of the virtuous look like a vice (Smith 1994, pp. 74 – 75). In what follows I offer an externalist explanatory account of a reliable connection between moral judgement and motivation. It is worth noting straight away that Smith's accusation is false. He wrongly thinks that externalism is condemned to explaining the motivation of good and virtuous persons as springing from natural features and non-moral contents – such as sympathy or compassion.

The task is to explain how virtuous persons are reliably motivated by *de re* desires from the externalist perspective. A distinction is, however, in order here. On the one hand, consider a person who is motivated by desire *de dicto* to do what he judges or believes to be right. *De dicto* desire "has a content that involves the concept of rightness; the concept of rightness figures as a part of the intentional content of [agent's] desire" (Strandberg 2007, p.4). On this view, agents perform actions because it is right: 'I want to do this because it is right'. On the other hand, "if a person is motivated by desire *de re* to do what she judges to be right, her desire does not have a content that involves the concept of rightness" (Strandberg 2007). Here an agent performs an action not because it is right, but because he desires or wants to do it. For example, Peter can say, 'I want to save drowning immigrants' or 'I desire to promote the people's well-being', I want to be more sympathetic or a loving father.

However, is it true that motivational externalism is condemned to *de dicto* explanation as Smith claims? I don't think so. It is possible, on the externalist construal, to account for a reliable connection in terms of *de re* desires. A person who judges that an action is morally right is motivated because he cares about morality and desires to keep faith with his moral-self. It is moral desire in tandem with moral care that gives him the motives to act on his moral judgements (that have the features that make actions morally right). In other words, his moral judgement alone cannot move him; and the desire in question is not *de dicto* but desire *de re*. He does not perform the action simply because it is right; rather, it is because he cares about it. This is because caring about morality and about one's self entails a higher-order desire (or care) to be moved not only de dicto by 'what is right', but also 'what is right' in the situation at hand. For example, for Ann to fully realize her moral self when it comes to keeping promises, she would need to develop a (caring) desire de re of keeping promises. That is, her motivating reason has to be because she gave her word, otherwise she would fail to realize her moral self fully. In other words, caring about morality entails caring about what morality entails as specific right actions.

Furthermore, there are also cases whereby an agent might judge that an action is morally right and she believes that the action in question has the right-making features. But if the strength of the moral desire to perform is weak, she will not be motivated. This explains some of the cases of motivational failures such as amoralism.

Recall the case of Peter. Assuming that it is morally required to save lives when certain persons are in danger; and Peter judges that the lives of the immigrants are in danger. However, assuming the moral desire to save them is weak; he would be not motivated accordingly. The desire to save the drowning immigrants needs to stem out of the fact that he (morally) cares or desires, and the action in question has the required moral features. But, if the moral desire or care (to realise one's moral self fully) is weak, the agent would remain unmoved. In other words, it is moral desire that links the moral-self, motivation and action. Further, the amoralist, for example, although he might have a strong self-understanding and grasps the features which make action morally right, but

he does not care so much about the desirability of his desire to be moral. In other words, he does not want the desire to be part of his will (Frankfurt 1971).

More so, recall Frankfurt's idea of first-order desires and second-order desires. Utilizing this model, I argue that the externalist account of the connection between moral judgement and motivation is not *de dicto* but *de re*. Consider an agent, Juliet who has a first-order (moral) desire, namely, to help someone in need. Notice that the concept of rightness does not figure in her desire. Nonetheless, we cannot qualify Juliet's desire as a *de re* desire, because 'helping someone in need' might still fall under a certain category of what is being said – *de dicto*. Suppose Juliet's friend Laura asks her: "But why do you want to help someone in need?" She might most likely reply thus: "Well, I want to act in a way that could be rightly described as helping someone". Precisely, Juliet's action is not strictly *about the person being helped* or *about helping someone in particular*. Thus, it is not a *de re* desire, but a *de dicto* desire. Consider another example, Laura desires to help street children (the desire in question is an independent moral desire). It is possible that her desire actually brings her to help the street children in Congo, because first-order desires can be effective desires (Frankfurt 1998, p.14). Nevertheless, such desires are not essentially part of the structure of her will. Alternatively, for her desires to be sufficient to motivate her to act accordingly, she must, among other things, possess some other desires that are not yet satisfied; and importantly, she must want such desires to be part of her will. In other words, Laura's desire must be second-order volition. She has to desire to desire to help the street children in Congo. Unlike Juliet's case, Laura's action is not desire *de dicto,* but *de re* desire. Precisely put, unless our deeper value provides second-order endorsement for the first-order desires, it seems 'wanton'. Alternatively, a second-order desire volition, if it has to motivate, needs to be consistent with the *de re* nature of the first-order desire. For first-order desire to enter into an agent's motivational stance, it has to possess *de re* nature, and not just about what is being said as in the case of Juliet above. It has to be about (certain) situations or persons in question. In the case of Laura, her first-order *de re* desire is about street children in Congo. Nevertheless, possessing a mere second-order desire does not fully motivate to act if it is not aligned with the *de re* nature of the first-order desire. Laura can desire 'helping street children' to be part of her

will, but this desire falls short if it is not realised in certain concrete situation – in her case, helping street children in Congo (of course, it can be street children anywhere). Also, this is the case with *de dicto* desire to act according to one's moral self-understanding. Recall Ann's example above, although 'promise keeping' is central to her moral self-understanding, her motivation would fall short if it is not consistent with her first-order *de re* desire, for example, keeping to her promise of driving her grandma to the dentist for a check-up.

More so, there are further cases where agents have a sense of moral-self coupled with desires to be good – *de dicto,* namely, let me do this because it is right or good. The content of such desires is not moral. On this level, agents act based on non-moral considerations. Although we might regard their actions as good (for example in cases where they hit the target randomly), their motives are not moral. Zangwill rightly puts it as follows: "Non-moral motivations may make us good but they cannot (by themselves) motivate us to act on our moral beliefs. For that we need a moral desire". (Zangwill 2003, p. 149) This is not moral motivation of virtuous persons, because the desires of the virtuous are moral. That is, they are rooted in the features that make actions morally right – features which are, in turn, in line with the moral ideals and values of the moral-self.

However, non-moral concerns can be moralized via moral education and moral self-creation. Metaphorically, moral education re-directs the agents from eating anything 'because it is fruit' to eating apple 'because it has the apple-making features'. Moral education or re-education also works for moral agents that relapsed from their original moral care and desires to performing action based on *de dicto* desires. They can regain their relevant desire *de re* via what Kristjánsson (2013) termed Aristotelian 'second-best tack', the view that those who were well raised, but later distanced themselves from moral values and ideals can through conscious self-determination reanimate their moral vigour. In other words, the moral self-conception can be re-created, and the desire to be moral will be re-integrated in the agent's will, such that the agent's actions not only capture the relevant moral features, but also spring from the moral self (where the agent possesses the desire to will all the way to actions).

7.7. Conclusion

The content of any moral desire is to do morally preferable things, things which have objective moral features and are in line with the moral identity of a moral person. The agent believes that the actions are not only morally preferable than their alternatives, but also they are at the core of the agent's moral-self. That is, he cares very much about morality. In other words, he wills the desires all the way to action because the action in question has moral features that make actions right; and he, in turn, cares about them.

However, one might object that it is the *moral-self or the desire to be faithful to the moral-self that determines the truth of the moral judgements.* The truth of moral judgements and consequent actions do not depend on the moral desires or moral-self (Zangwill 2003, p. 146). We do not perform moral actions because they are merely *ad rem* to what we feel or yearn to do. Rather, it is because the action in question has the morally right-making features and those features are mirrored on the moral-self. These features are, metaphorically, incarnated in the moral-self, such that they become an integral or holistic part of the moral person. Zangwill rightly points out that moral desires do not make agents virtuous, they only motivate them (Zangwill 2003). In sum, moral motivation is mediated through the will to maintain and express one's moral identity (Nisan 1996, p. 78). Moral agents perform actions that are in unity with their moral identity, not just to avoid self-inconsistency, but also to express their moral signatures. It is via the positive affirmation of the self that others identify the characterizing features that make moral persons "who they are", "what they stand for morally", and "where they stand morally".

CONCLUSION

To answer the question pertaining to the practical aspect of morality: "In what sense, if at all, can objective moral judgement motivate agents to act morally? I addressed core theoretical issues. I argued that our straightforward intuition about morality is that when we engage in moral discourse, we talk about something, which can be true or false. We understand moral discourse as propositionally laden. We encountered that philosophers, even the expressivists, are increasingly accepting the propositional features of moral discourse. However, I argued that the propositional surface alone as – quasi-realism and minimalism hold – is insufficient to secure our intuition about the theoretical aspect of morality. Precisely, I argued that, unlike revisionism which fails in advance in providing the semantic and logical features of moral discourse, the hybrid expressivist minimalistic construal of truth-value does not capture the core idea of moral discourse, namely, moral objectivity. Its internal reading technique makes moral discourse attitude-dependent. The resources to make moral discourse mind-independent are not available to hybrid expressivism. I argued that moral realism seems to do justice to moral discourse. Consequently, I argued for the core features of moral realism, which not only served to separate moral realism from other forms of realism (e.g. quasi-realism), but also to show that moral realism entails cognitivism and not vice versa. Its metaphysics and robust construal of truth-aptness enable moral realism to account for the semantic and logical features just like any other domain of ordinary discourse. However, I do not claim that these core features settle the debate on how we should conceive the ontological commitments of moral realism. Nevertheless, one of the advantages of this map-off model is that it allows us to plug in an 'ontological-commitment' relevant for understanding the practical aspect of morality.

Having taken a metaphysical stand, namely that moral facts are non-reductively naturalistic, I argued that externalism has better resources for explaining the psychology of moral motivation, given that the realist internalist solution is insufficient. Given its social practical dimension, I showed that morality is vital in personality development as well as in shaping our self-understanding. When an agent cares about morality, she

desires its realisation. The degree of agents' moral care explains the varied motivational manifestations ranging from motivations in virtuous persons and indifferent persons to amoralists. I argued that when an agent desires morality to be an essential part of 'who she is', it triggers off the commitment to be faithful to her moral self. In other words, the fully realisation of self occurs when an agent is faithful to the core moral values – moral signatures – she cares about; and the moral desire to see those moral values full realised, in turn, motivates her to act morally.

However, I suspect the reader will be disappointed, if he understands the externalist account defended here as a definitive solution to moral motivational problem. My main objective is rather to offer a plausible way of accounting for how an objective moral judgement motivates agents to act morally; and to urge more research on the role of moral care and moral self in the internalist-externalist debate. It is hoped that such an exercise will enable us to offer solutions to complex social and environmental issues such as the motivational gap in climate change inaction.

For now, whether or not we believe that externalism better answers the moral motivational problem; and whether or not we believe that moral care and moral desire to be faithful to our moral-selves make up our motivational profiles as social agents, at least most of us believe that they are nevertheless very crucial to moral agency, and as such, we can pull our efforts together in offering a vivid picture of what drives our motivation to act morally. Let the work begin!

REFERNCES

Allport, Gordon W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Allport, Gordon W. & Harrington, Milton (1938). Personality: A Psychological Interpretation. *Ethics* 49 (1):105-107.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (IV) Washington, DC.

Anscombe, G. E. M. (1957). *Intention.* Harvard University Press.

Aquino, K., & Reed, A. II. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423-1440.

Aquino, K., Freeman, D., Reed, A., & Lim, V.K.G. (2009). Testing a social-cognitive model of moral behavior: The interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology (97):* 123-141.

Ayer, Alfred Jules (1952). *Language, Truth and Logic.* Courier Corporation.

Baier, Annette (1991). *A Progress of Sentiments: Reflections on Hume's Treatise.* Harvard University Press.

Baker, Lynne Rudder (2001). Philosophy in Mediis Rebus. *Metaphilosophy* 32:378-394.

Bambrough, Renford (1979). *Moral Skepticism and Moral Knowledge.* Routledge + Kegan Paul.

Bechara, Antoine, Damasio, Hanna, & Damasio, Antonio, R. (2000). Emotion, Decision Making and the Orbitofrontal Cortex. *Cerebral Cortex* 10 (3):295–307.

Benn, S. (1985). Wickedness. *Ethics 95* (3): 795–810.

Berlin, Isaiah (2001). *The Power of Ideas.* Henry Hardy (ed.) London: Chatto & Windus.

Björklund, Fredrik; Björnsson, Gunnar; Eriksson, John; Francén Olinder, Ragnar & Strandberg, Caj (2012). Recent Work on Motivational Internalism. *Analysis* 72 (1):124-137.

Björnsson, Gunnar & Francén Olinder, Ragnar (2013). Internalists Beware—we Might all be Amoralists! *Australasian Journal of Philosophy* 91 (1):1 - 14.

Björnsson, Gunnar; Strandberg, Caj; Francén Olinder, Ragnar; Eriksson, John & Björklund, Fredrik (2015). Motivational Internalism: Contemporary Debates. In Gunnar Björnsson, Caj Strandberg, Ragnar Francén Olinder, John Eriksson & Fredrik Björklund (eds.), *Motivational Internalism*. Oxford University Press 1–20.

Blackburn, Simon (1981). Reply: Rule-following and moral realism. In S. Holtzman & Christopher M. Leich (eds.), *Wittgenstein: To Follow a Rule*. Routledge 163--87.

Blackburn, Simon (1984). *Spreading the Word*. Clarendon Press.

Blasi, Augusto (1984). Moral identity: Its role in moral functioning. In W.M. Kurtines & J. J. Gewirtz (eds.), *Morality, moral behavior and moral development*. New York: John Wiley and Sons. pp. 128 – 139.

Blasi, Augusto (1995). Moral understanding and the moral personality: The process of moral integration. In W. M. Kurtines & J. J. Gewirtz (eds.), *Moral development: An introduction*. Boston: Allyn & Bacon. pp. 229 –253.

Blasi, Augusto (2004). Moral functioning: Moral understanding and personality. In D. K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity.* Mahwah: Erlbaum. pp. 335–347.

Blasi, Augusto. (2004). Neither personality nor cognition: An alternative approach to the nature of the self. In C. Lightfoot, C. Lalonde & M. Chandler (eds.), *Changing conceptions of psychological life* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum Associates.

Blasi, Augusto (2005). Moral character: A psychological approach. In D. K. Lapsley & F.C. Power (eds.), *Character psychology and character education.* University of Notre Dame Press. pp. 18-35.

Blok, Sergey, Newman, George, & Rips, Lance J. (2005). Individuals and their concepts. In W.-K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in Honor of Douglas L. Medin* (pp. 127–149). Washington, DC: American Psychological Association.

Boyd, Richard (1988). How to be a Moral Realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Cornell University Press 181-228.

Bricke, John (1996). *Mind and Morality: An Examination of Hume's Moral Psychology*. Oxford University Press.

Brink, David O. (1986). Externalist moral realism. *Southern Journal of Philosophy* 24 (S1):23-41.

Brink, David Owen (1989). *Moral Realism and the Foundations of Ethics*. Cambridge University Press.

Bromwich, Danielle Nicole (2008). *Belief Internalism*. Retrieved from TSpace Repository/School of Graduate Studies – Theses. 2009-01-20T19:35:44Z.

Bromwich, Danielle (2010). Clearing conceptual space for cognitivist motivational internalism. *Philosophical Studies* 148 (3):343 - 367.

Bromwich, Danielle (2013). Motivational Internalism and the Challenge of Amoralism. *European Journal of Philosophy* 24 (1): n/a-n/a.

Brown, C. (1988). Is Hume an internalist? *Journal of the History of Philosophy* (26): 69–87.

Carr, David (2001). Moral and personal identity. *International Journal of Education and Religion*, II (1), 79-97.

Cassam, Quassim (2007). The possibility of knowledge. *Grazer Philosophische Studien* 74 (1):125-141.

Cervone, D. & Tripathi, R. (2009). The Moral Functioning of the Person as a Whole: On Moral Psychology and Personality Science. In D. Narvaez & D. Lapsley (eds.), *Personality, Identity, and Character: Explorations in Moral Psychology*. Cambridge University Press. pp. 30-51.

Cholbi, Michael (2006). Belief attribution and the falsification of motive internalism. *Philosophical Psychology* 19 (5):607 – 616.

Cholbi, Michael (2011). Depression, listlessness, and moral motivation. *Ratio* 24 (1):28-45.

Cohon, Rachel (1993). Internalism about reasons for action. *Pacific Philosophical Quarterly* 74 (4):265-288.

Cohon, Rachel (1994). On an Unorthodox Account of Hume's Moral Psychology. *Hume Studies* 20 (2):179-194

Cohon, Rachel (1988). Hume and Humeanism in ethics. *Pacific Philosophical Quarterly,* 69, 99–116.

Colby, Anne & Kohlberg, Lawrence (Eds.) (1987). *The Measurement of Moral Judgment* (Vols. 1 and 2). New York: Cambridge University Press.

Colby, A. & Damon, W. (1992). *Some do care: Contemporary lives of moral commitment.* New York: Free Press.

Coleman, Dorothy (1992). Hume's Internalism. *Hume Studies* 18 (2):331-347.

Coleman, Mary Clayton (2008). Directions of fit and the Humean theory of motivation. *Australasian Journal of Philosophy* 86 (1):127 – 139.

Crowder, G. (2002). *Liberalism and Value Pluralism*, Continuum, London.

Cuneo, Terence (2007). Recent Faces of Moral Nonnaturalism. *Philosophy Compass* 2 (6):850-879.

Cuneo, Terence (2007). *The Normative Web: An Argument for Moral Realism*. Oxford University Press.

Cuneo, Terence & Shafer-Landau, Russ (2014). The moral fixed points: new directions for moral nonnaturalism. *Philosophical Studies* 171 (3):399-443.

Damasio, Antonio R. (1995). On Some Functions of the Human Prefrontal Cortex. *Annals of the New York Academy of Sciences*, 769: 241–252.

Damasio Antonio R., Tranel, D. & Damasio, Hanna (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research* 41(2):81–94.

Dancy, Jonathan (1993). *Moral Reasons*. Blackwell.

Dancy, Jonathan (2000). *Practical Reality*. Oxford University Press.

Dancy, Jonathan (ed.) (2000). *Normativity*. Blackwell Publishers.

Darwall, Stephen L. (1983). *Impartial Reason*. Cornell University Press.

Darwall, Stephen L. (1992). Internalism and agency. *Philosophical Perspectives* 6:155-174.

De Sousa, Ronald B. (1987). *The Rationality of Emotion*. MIT Press.

Dworkin, Ronald (1996). Objectivity and Truth: You'd Better Believe It. *Philosophy and Public Affairs* 25 (2):87-139.

Enoch, David (2011). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford University Press.

Erikson, E. (1964). *Insight and responsibility*. New York: Norton.

Falk, W. D. (1945). Obligation and Rightness. *Philosophy* 20 (76):129 - 147.

Falk, W. D. (1986). *Ought, Reasons, and Morality: The Collected Papers of W.D. Falk*. Cornell University Press.

Fine, Kit (1994). Essence and modality. *Philosophical Perspectives* 8:1-16.

Fitzpatrick, William (2008). Robust ethical realism, non-naturalism, and normativity. *Oxford Studies in Metaethics* 3:159-205.

Frijda, Nico H. (2009). Emotion Experience and its Varieties. *Emotion Review* 1 (3):264-271.

Foot, Philippa (1972). Morality as a system of hypothetical imperatives. *Philosophical Review* 81 (3):305-316.

Foot, Philippa (1978). *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford University Press.

Frankena, W. K. (1958). Obligation and motivation in recent moral philosophy. In A. I. Melden (ed.), *Essays in Moral Philosophy*. University of Washington Press

Frankfurt, Harry G. (2006). *Taking Ourselves Seriously & Getting It Right*. Stanford University Press.

Frankfurt, Harry G. (1998). *The Importance of What We Care About.* Cambridge University Press.

Frankfurt, Harry G. (1999). Necessity, Volition and Love. Cambridge University Press.

Frankfurt, Harry G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy* 68 (1):5-20.

Frankfurt, Harry (2004). *The Reasons for Love.* Princeton University Press.

Gelman, S. A., Heyman, G. D., & Legare, C. H. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child Development, 78*, 757-774.

Gewirth, Alan (1960). Meta-ethics and normative ethics. *Mind* 69 (274):187-205.

Gibbard, Allan (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.

Gibbard, Allan (2003). *Thinking How to Live*. Harvard University Press.

Goodwin, Geoffrey P., Piazza, Jared & Rozin, Paul (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, Vol 106(1), Jan 2014, 148-168.

Hall, Everett W (1947). Stevenson on Disagreement in Attitude. *Ethics* 58 (1): 51-56.

Hampton, Jean (1995). Does Hume Have an Instrumental Conception of Practical Reason? *Hume Studies* 21 (1):57-74.

Hardy, Sam A. & Carlo, Gustavo (2005). Religiosity and prosocial behaviours in adolescence: the mediating role of prosocial values. *Journal of Moral Education* 34 (2):231-249.

Hare, Robert (1993). *Without Conscience: The Disturbing World of the Psychopaths among Us*. New York: Guilford Press.

Hare, R. M. (1952). *The Language of Morals*. Oxford Clarendon Press.

Haslam, S. A., Jetten, J., O'Brien, A., & Jacobs, E. (2004). Social identity, social influence, and reactions to potentially stressful tasks: Support for the self-categorization model of stress. *Stress and Health,* 20, 3–9.

Hobbes, Thomas (1651). *Leviathan*. Dover Publications.

Hoffmann, Tobias (2008). Henry of Ghent's Voluntarist Account of Weakness of Will. In *Weakness of Will from Plato to the Present*. Catholic University of America Press.

Horgan, Terry & Timmons, Mark (2006). Morality without Moral Facts. In James Lawrence Dreier (ed.), *Contemporary Debates in Moral Theory*. Blackwell Pub. 6--220.

Hume, David (1739/2000). *A Treatise of Human Nature*. Oxford University Press.

Hume, David (1777). *An Enquiry Concerning the Principles of Morals*. Prometheus Books.

Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press.

Jaworska, Agnieszka (2007). Caring and full moral standing. *Ethics* 117 (3):460-497.

Johnson, Robert N. (1993). *Kant's Theory of Moral Worth*. Dissertation, The University of North Carolina at Chapel Hill.

Joyce, Richard (2001). *The Myth of Morality*. Cambridge University Press.

Joyce, Richard (2010). Expressivism, motivation internalism, and Hume. In Charles R. Pigden (ed.), *Hume on Motivation and Virtue*. Palgrave Macmillan.

Kennett, Jeanette (2006). Do psychopaths really threaten moral rationalism? *Philosophical Explorations* 9 (1):69 – 82.

Kohlberg, Lawrence (1981). *The Philosophy of Moral Development Moral Stages and the Idea of Justice.* Harper & Row.

Kohlberg, Lawrence (1969). Stage and sequence: The cognitive development approach to socialization. In D. A. Goslin (ed.), *Handbook of socialization theory* Rand McNally.pp. 347-480.

Korsgaard, Christine M. (2008). *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford University Press.

Korsgaard, Christine M. (1996). *The Sources of Normativity*. Cambridge University Press.

Kramer, Matthew H. (2009). *Moral Realism as a Moral Doctrine*. Wiley-Blackwell.

Kristjánsson, Kristján (2013). Aristotelian motivational externalism. *Philosophical Studies* 164 (2):419-442.

Kühler, Michael (2012). Who Am I to Uphold Unrealizable Normative Claims? In Michael Kühler & Nadja Jelinek (eds.), *Autonomy and the Self*. Springer. pp. 191--209.

Lapsley, Daniel K. (2007). Moral Self-Identity as the Aim of Education, in edited by Nucci, L. and Narvaez, D. (Eds.), *Handbook of Moral and Character Education*. Lawrence Erlbaum Associates, Mahwah, NJ. (pp. 30–52).

Lapsley, Daniel K. & Hill, Patrick L. (2009). The development of the moral personality. In Darcia Narvaez & Daniel Lapsley (eds.), *Personality, Identity, and Character*. Cambridge University Press. pp. 185 - 213.

Lapsley, Daniel K., & Narvaez, Darcia (2004). A social-cognitive approach to the moral personality. In D.K. Lapsley & D. Narvaez (eds.), *Moral development, self and identity* Mahwah: Lawrence Erlbaum. pp. 189-212.

Leben, Derek & Wilckens, Kristine (2014). Pushing the Intuitions behind Moral Internalism. *Philosophical Psychology* 28 (4):510-528.

Lewis, David (1983). *Philosophical Papers.* Oxford: Oxford University Press.

Lykken, David (1995). *The Antisocial Personalities.* Hillsdale: Lawrence Erlbaum

Macintyre, A. (1984). After Virtue. *Tijdschrift Voor Filosofie* 46 (1):169-171.

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Penguin.

Mackie, J. L. (1980). *Hume's Moral Theory*. Routledge & K. Paul.

Maibom, Heidi Lene (2005). Moral unreason: The case of psychopathy. *Mind and Language* 20 (2):237-57.

McAdams, D. (2009). The Moral Personality. In D. Narvaez & D.K. Lapsley (eds.), *Personality, identity and character: Explorations in moral psychology*. New York: Cambridge University Press. pp. 11-29.

192

McDowell, John (1978). Are Moral Requirements Hypothetical Imperatives? *Proceedings of the Aristotelian Society* 52:13-29+31-42.

McDowell, John (1979). Virtue and Reason. *The Monist* 62 (3):331-350.

McDowell, John (1981). Anti-realism and the epistemology of understanding. In Herman Parret & Jacques Bouveresse (eds.), *Meaning and Understanding*. W. De Gruyter 225--248.

McDowell, John (1981). Following a Rule and Ethics. In S. Holtzman & Christopher M. Leich (eds.), *Wittgenstein: To Follow a Rule*. Routledge 141--162.

McNaughton, David (1988). *Moral Vision: An Introduction to Ethics*. B. Blackwell.

Mele, Alfred R. (1996). Internalist moral cognitivism and listlessness. *Ethics* 106 (4):727-753.

Meyers, Diana Tietjens (2014). Corporeal selfhood, self-interpretation, and narrative selfhood. *Philosophical Explorations* 17 (2):141-153.

Miller, Alexander (1996). An objection to Smith's argument for internalism. *Analysis* 56 (3):169–174.

Miller, Alexander (1998). Emotivism and the verification principle. *Proceedings of the Aristotelian Society* 98 (2):103–124.

Miller, Alexander (2003). *An Introduction to Contemporary Metaethics*. Distributed in the USA by Blackwell Publishers.

Miller, Alexander (2006). Realism and Antirealism. In Barry C. Smith (ed.), *The Oxford Handbook of Philosophy of Language*. Oxford University Press 983.

Miller, Christian (2013). The Euthyphro Dilemma. In *Blackwell International Encyclopedia of Ethics*. Blackwell 1-7.

Milo, Ronald D. (1981). Moral Indifference. *The Monist* 64 (3):373-393.

Moore, G. E. (1903). *Principia Ethica*. Dover Publications.

Murdoch, Iris (2013). *The Sovereignty of Good*. Routledge.

Nagel, Thomas (1970). *The Possibility of Altruism*. Oxford Clarendon Press.

Nagel, Thomas (1979). *The Limits of Objectivity*. The Tanner Lecture on Human Values, Brasenose College, Oxford University.

Narvaez, Darcia & Lapsley, Daniel K. (2009). Moral identity, moral functioning and the development of moral character. In D. M. Bartels, Ch. W. Bauman, L. J. Skitka, & D. L. Medin, (eds.), *The Psychology of Learning and Motivation*. Burlington: Academic Press. pp. 237-274.

Narvaez, Darcia & Lapsley, Daniel K. (2009). Moral personality: themes, questions, futures. In Darcia Narvaez & Daniel Lapsley (eds.), *Personality, Identity, and Character*. Cambridge University Press.

Nichols, Shaun (2002). How Psychopaths Threaten Moral Rationalism. *The Monist* 85 (2):285-303.

Nichols, Shaun (2004). After objectivity: An empirical study of moral judgment. *Philosophical Psychology* 17 (1):3 – 26.

Nichols, Shaun (2006). Folk intuitions on free will. *Journal of Cognition and Culture* 6 (1):57-86.

Nichols, Shaun & Bruno, Michael (2010). Intuitions about personal identity: An empirical study. *Philosophical Psychology* 23 (3):293-312.

Nisan, Mordecai (1996). Personal Identity and Education for the Desirable. *Journal of Moral Education* 25 (1):75-83.

Noggle, Robert (1997). The nature of motivation (and why it matters less to ethics than one might think). *Philosophical Studies* 87 (1):87-111.

Oliner, Samuel P. & Oliner Pearl M. (1988). *The Altruistic Personality*. New York: The Free Press.

Pathak, R., Sharma R.C., Parvan, U.C., Gupta, B. P., Ojha, R.K. & Goel, N. (2011).Behavioural and Emotional Problems in School going Adolescents. *The Australasian Medical Journal*. 4(1):15 - 21.

Parfit, Derek (1971). Personal identity. *Philosophical Review* (80):3-27.

Parfit, Derek (1971). On the importance of self-identity. *Journal of Philosophy* 68: 683-90.

Parfit, Derek (1984). *Reasons and Persons*. Oxford University Press.

Perry, John (2002). *Identity, Personal Identity, and the Self*. Monograph Collection (Matt - Pseudo).

Pettit, Philip (1987). Humeans, anti-Humeans, and motivation. *Mind* 96 (384):530-533.

Piaget, Jean (1933). The Moral Judgement of the Child. *Philosophy* 8 (31):373-374.

Pigden, Charles R. (2007). Hume, motivation and "the moral problem". *Rivista di Storia Della Filosofia* 62 (3):199-221.

Pigden, Charles (2009). Introduction to Hume on Motivation and Virtue. In *Hume on Motivation and Virtue*. 1-29.

Platts, Mark de Bretton (1980). Moral Reality and the End of Desire. In Mark Platts (ed.), *Reference, Truth, and Reality*. Routledge and Kegan Paul

Platt des, Mark Bretton (1979). *Ways of Meaning: An Introduction to a Philosophy of Language*. Routledge & K. Paul.

Prichard, James Cowles (1842). *Forms of Insanity*: Aubrey Lewis, Psychopathic personality: a most elusive catergory (1974) 4 Psychological Medicine 133.

Putnam, Hilary (2004). *Ethics without Ontology*. Harvard University Press.

Radcliffe, Elizabeth S. (1999). Hume on the Generation of Motives: Why Beliefs Alone Never Motivate. *Hume Studies* 25 (1-2):101-122.

Railton, Peter (1986). Moral realism. *Philosophical Review* 95 (2):163-207.

Railton, Peter (1986). Facts and values. *Philosophical Topics* 14 (2):5-31.

Railton, Peter (1989). Naturalism and Prescriptivity. *Social Philosophy and Policy* 7 (1):151.

Railton, Peter (2003). Facts, Values, and Norms. *Philosophical Studies* 126 (3):449-462.

Railton, Peter (2006). Humean theory of practical rationality. In David Copp (ed.), *The Oxford Handbook of Ethical Theory*. Oxford University Press 265--81.

Railton, Peter (2006). Moral factualism. In James Lawrence Dreier (ed.), *Contemporary Debates in Moral Theory*. Blackwell Pub. 6--201.

Rawls, John (2009). A theory of justice. In Steven M. Cahn (ed.), *Philosophy and Rhetoric*. Oxford University Press 133-135.

Rosati, Connie S. (2006). Moral motivation. *Stanford Encyclopedia of Philosophy*.

Roskies, Adina L. (2003). Are ethical judgments intrinsically motivational? Lessons from "acquired sociopathy". *Philosophical Psychology* 16 (1):51 – 66.

Roskies, Adina L. (2006). Neuroscientific challenges to free will and responsibility. *Trends in Cognitive Sciences* 10 (9):419-423.

Rudder Baker, Lynne (2001). Philosophy in Mediis Rebus. *Metaphilosophy* 32 (4):378-394.

Sayre-McCord, Geoffrey (ed.) (1988). *Essays on Moral Realism*. Cornell University Press.

Scanlon, Thomas (1998). *What We Owe to Each Other*. Belknap Press of Harvard University Press.

Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford University Press.

Schroeder, Mark Andrew (2007). *Slaves of the Passions*. Oxford University Press.

Schroeder, Mark (2007). The Humean theory of reasons. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Vol. 2*. Oxford University Press 195 - 219.

Schroeder, Mark (2010). *Noncognitivism in Ethics*. Routledge.

Schaber, Peter (1997) Moralischer Realismus. Freiburg/München.

Shafer-Landau, Russ (1998). Moral judgement and Moral Motivation. *Philosophical Quarterly* 48 (192):353-358.

Shafer-Landau, Russ (2000). A defense of motivational externalism. *Philosophical Studies* 97 (3):267-291.

Shafer-Landau, Russ (2003). *Moral Realism: A Defence*. Oxford University Press.

Shafer-Landau, Russ (ed.) (2008). *Oxford Studies in Metaethics, Volume 3*. Oxford University Press.

Shao, R., Aquino, K., & Freeman, D. (2008). Beyond Moral Reasoning: A Review of Moral Identity Research and Its Implications for Business Ethics. *Business Ethics Quarterly, 18*(4): 513-540.

Sher, George (ed.) (2012). *Ethics: Essential Readings in Moral Theory*. Routledge.

Shoemaker, Sydney (1959). Personal identity and memory. *Journal of Philosophy* (56):868-902.

Sinclair, Neil (2008). Free Thinking for Expressivists. *Philosophical Papers* 37 (2):263-287.

Skarsaune, Knut Olav (2015). How to Be a Moral Platonist. *Oxford Studies in Metaethics* (10).

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? Journal of Personality and Social Psychology, 88(6): 895–917.

Smith, Adam (1790). *The Theory of Moral Sentiments*. Dover Publications.

Smith, Michael (1992). Valuing: Desiring or Believing? In K. Lennon & D. Charles (eds.), *Reduction, Explanation, and Realism*. Oxford University Press 323--60.

Smith, Michael (1987). The Humean theory of motivation. *Mind* 96 (381):36-61.

Smith, Michael (1994). *The Moral Problem*. Blackwell.

Smith, Michael (1996). The argument for internalism: Reply to Miller. *Analysis* 56 (3):175–184.

Smith, Michael (2000). Moral realism. In Hugh LaFollette (ed.), *The Blackwell Guide to Ethical Theory*. Blackwell Publishers 15--37.

Smith, Deborah C. (2001). Moral Realism, Skepticism and Anti-Realism: A critical analysis of the criteria for moral realism. *Disputatio* 11:1 - 10.

Stevenson, Charles Leslie (1937). The emotive meaning of ethical terms. *Mind* 46 (181):14-31.

Stevenson, Charles Leslie (1944). *Ethics and Language*. London: Oxford Press

Stevenson, Charles L. (1998). The nature of ethical disagreement. In James Rachels (ed.), *Ethical Theory: The Question of Objectivity*. Oxford University Press.

Stocker, Michael (1979). Desiring the bad: An essay in moral psychology. *Journal of Philosophy* 76 (12):738-753.

Strandberg, Caj & Björklund, Fredrik (2013). Is Moral Internalism Supported by Folk Intuitions? *Philosophical Psychology* 26 (3):319-335.

Strandberg, Caj (2013). An Internalist Dilemma—and an Externalist Solution. *Journal of Moral Philosophy* 10 (1):25-51.

Strandberg, Caj (2000). Aristotle's internalism in the nichomachean ethics. *Journal of Value Inquiry* 34 (1):71-87.

Strandberg, Caj (2013). Cognitivism. In Hugh LaFollette (ed.), *International Encyclopedia of Ethics*. Blackwell.

Strandberg, Caj (2004). *Moral Reality. A Defence of Moral Realism*. Lund University.

Strandberg, Caj (2007). Externalism and the content of moral motivation. *Philosophia* 35 (2):249-260.

Street, Sharon (2011). Mind-Independence Without the Mystery: Why Quasi-Realists Can't Have It Both Ways. *Oxford Studies in Metaethics* (6):1-32.

Strohminger, Nina & Nichols, Shaun (2014). The essential moral self. *Cognition* 131 (1):159-171.

Stroud, Barry (1977). *Hume*. Routledge.

Sturgeon, Nicholas L. (1986). What difference does it make whether moral realism is true? *Southern Journal of Philosophy* (24):115-141.

Sturgeon, Nicholas L. (1998). Naturalism in Ethics. In Edward Craig (ed.), *Routledge Encyclopedia of Philosophy*. Routledge

Sturgeon, Nicholas L. (2002). Ethical Intuitionism and Ethical Naturalism. In Phillip Stratton-Lake (ed.), *Ethical Intuitionism: Re-evaluations*. Oxford University Press.

Sturgeon, Nicholas L. (2006). Ethical Naturalism. In David Copp (ed.), *The Oxford Handbook of Ethical Theory*. Oxford University Press.

Sugarman, J. (2005). Persons and moral agency. *Theory and Psychology (15):* 793–811.

Sullivan, Dennis M. (2003). The Conception View of Personhood. *Ethics and Medicine* (19): 11-33.

Svavarsdóttir, Sigrun (1999). Moral cognitivism and motivation. *Philosophical Review, 108*, 161–219

Sytsma, Sharon E. (1995). Ethical internalism and moral indifference. *Journal of Value Inquiry* 29 (2):193-201.

Taylor, Charles (1989). *Sources of the Self: The Making of the Modern Identity*. Harvard University Press.

Tietjens Meyers, Diana Tietjens (2012). Psychocorporeal Selfhood, Practical Intelligence, and Adaptive Autonomy. In Michael Kuhler & Najda Jelinek (eds.), *Autonomy and the Self*. springer. pp. 147–172.

Timmons, Mark & Baiasu, Sorin (eds.) (2013). *Kant on Practical Justification: Interpretative Essays*. Oxford University Press.
Tresan, Jon (2009). Metaethical internalism: Another neglected distinction. *Journal of Ethics* 13 (1):51 - 72.

van Ackeren Marcel, & Kühler Michael, (eds.) (2016). *The Limits of Moral Obligation*. Routledge.

van Roojen, Mark (1995). Humean motivation and Humean rationality. *Philosophical Studies* 79 (1):37-57.

van Roojen, Mark (2002). Humean and anti-Humean internalism about moral judgements. *Philosophy and Phenomenological Research* 65 (1):26-49.

Vargas, Manuel & Nichols, Shaun (2007). Psychopaths and moral knowledge. *Philosophy, Psychiatry, and Psychology* 14 (2):157-162.

Velleman, David (2002). Idenitification and Identity. In Sarah Buss & Lee Overton (eds.*), Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge/MA: MIT Press. pp. 91 – 123.

Warenski, Lisa (2014). Defending Moral Mind-Independence: The Expressivist's Precarious Turn. *Philosophia* 42 (3):861-69.

Wedgwood, Ralph (1995). Theories of content and theories of motivation. *European Journal of Philosophy* 3 (3):273-288.

Weiss, H. M & Cropanzano, R. (2014). Affective Events Theory: A Theoretical Discussion of the Structure, Causes and Consequences of Affective Experiences at Work," in B. M. Staw and L. L. Cummings (eds.), *Research in Organizational Behavior,* 18 (Greenwich, CT: JAI Press, 1996), pp. 17–19.

Wielenberg, Erik J. (2014). *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism*. OUP Oxford.

Williams, Bernard Arthur Owen (1981). *Moral Luck: Philosophical Papers, 1973-1980*. Cambridge University Press.

Williams, Bernard (1981). Persons, Character, and Morality. In James Rachels (ed.), *Moral Luck*. Cambridge University Press.

Wittgenstein, Ludwig (1965). I: A lecture on ethics. *Philosophical Review* 74 (1):3-12.

Zangwill, Nick (1994). Moral mind-independence. *Australasian Journal of Philosophy* 72 (2):205-219.

Zangwill, Nick (2003). Externalist moral motivation. *American Philosophical Quarterly* 40 (2):143-154.
Zangwill, Nick (2008). The indifference argument. *Philosophical Studies* 138 (1):91 - 124.