

6-2014

Moral Motivation and the Authority of Morality: A Defense of Naturalist Moral Realism

Lily Eva Frank

Graduate Center, City University of New York

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

 Part of the [Ethics and Political Philosophy Commons](#)

Recommended Citation

Frank, Lily Eva, "Moral Motivation and the Authority of Morality: A Defense of Naturalist Moral Realism" (2014). *CUNY Academic Works*.

https://academicworks.cuny.edu/gc_etds/211

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.

**MORAL MOTIVATION AND THE AUTHORITY OF MORALITY: A DEFENSE OF
NATURALIST MORAL REALISM**

by

LILY EVA FRANK

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2014

© 2014

LILY EVA FRANK

All Rights Reserved

This manuscript has been read and accepted for the
Graduate Faculty in Philosophy in the satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Jesse Prinz

Date

Chair of Examining Committee

Iakovos Vasiliou

Date

Executive Officer

Steven Cahn

Rosamond Rhodes

Stefan Baumrin

Jesse Prinz

Hagop Sarkissian

THE CITY UNIVERSITY OF NEW YORK

Abstract
MORAL MOTIVATION AND THE AUTHORITY OF MORALITY: A DEFENSE OF
NATURALIST MORAL REALISM

by

Lily Eva Frank

Adviser: Professor Steven Cahn

Moral realism has been continuously accused of positing the existence of queer properties, facts, judgments, and beliefs. One of these queer features is supposed to be the normative force of morality—that is the way in which morality guides our actions. Critics of moral realism argue that nothing else in the world has this feature. This is a reason to doubt that moral facts and properties exist at all. This objection can be interpreted in at least two ways. One way to interpret it has to do with moral motivation, this is the internalism objection. The other has to do with the authority of morality. In this essay I defend naturalist moral realism against these two objections, the internalism objection and the authority objection.

I argue that the internalism objection and the authority objection are independent of each other. Whether and how morality motivates us to act does not bear on the place that morality should have in our lives and decision-making. We may have no motivation to do things that we should do, and we may be extremely motivated to do things we should not do. The conflation of these two objections is widespread in the literature and is the source of some of their apparent persuasiveness.

The internalism objection is an attack on the realist commitment to cognitivism—the view that moral judgments are beliefs. The objection claims that there is a necessary, essential, or inherent connection between making a moral judgment, and

being motivated to some extent to act on that moral judgment. This claim is internalism. But, the objection claims, beliefs, by themselves, cannot motivate an agent to act. This means that moral judgments cannot be beliefs. So cognitivism is false. The objection concludes that since moral realism is wedded to cognitivism, moral realism is also false.

I respond to this objection in two ways. First, I argue against the view that moral judgments necessarily, inherently, or essentially motivate. I argue against it based in part on the phenomena of amoralism and other forms of moral indifference, including mundane, everyday cases of occasional amoralism and diversity in the level of moral motivation. The morally indifferent seem to be able to make the same moral judgments that morally deferential people do, without having any corresponding motivation to act on those judgments. Thus they present a central counter-example to the internalist's claim that there is a necessary, essential, or inherent connection between moral judgment and motivation. I stress that moral indifference should not be understood as an exotic phenomenon found only in psychopathic serial killers or suicidally depressed. Instead, it is commonplace. Moral indifference might actually be the majority of our moral experience, rather than the exceptional case. I also argue against the internalist method of a priori conceptual analysis.

At the same time, I make a case for the opposite view, externalism, which is the view that moral judgments do not necessarily or inherently motivate, nor can they motivate by themselves. Instead moral judgments are only contingently connected with motivation. The specific form of externalism that I argue for is a pluralistic externalism, which I argue can meet the objections that are usually made against externalism better than any alternative form of externalism.

The authority objection to naturalist moral realism is that morality has a certain kind of authority over us and that naturalist moral realism precludes this kind of authority. Therefore, naturalist moral realism must be false. The authority of morality can be understood in a variety of ways. For example, the importance that moral demands have in directing our lives or the way in which moral reasons seem to override other reasons for action. The authority of morality is supposed to be a problem for naturalist moral realism because the realist identifies moral facts and properties with complex natural facts and properties. The authority objection asks: why should any set of natural facts or properties have authority over our behavior? In other words, the naturalist moral realist seems to lack a convincing response to this kind of moral skeptic. One particularly influential form of this argument is found in the work of Christine Korsgaard, who argues that as a metaphysical position, moral realism is ill equipped to account for the normativity of morality. The other is found in Derek Parfit, who argues that naturalist forms of realism cannot explain normativity.

I respond to the authority objection by defending a limited account of authority. Second, I argue that once properly understood, the authority of morality is no more a problem for naturalist moral realism as a metaethical theory than any other meta-ethical theory. Every metaethical position is faced with the difficult task of explaining this aspect of normativity and we have no reason to think this is a special problem for realism. Finally, I put forward a defensible version of naturalist moral realism, spelling out the commitment to objectivity and to naturalism.

Acknowledgments

I thank my adviser Steven Cahn, whose guidance, wisdom, and sense of tranquility were exactly what I needed during this stage of graduate school.

I thank my committee members Stefan Baumrin and Rosamond Rhodes. I am indebted to Stefan, whose incisive questions and acerbic wit truly made this a better project. I thank Rosamond, who has been my mentor throughout graduate school. In many ways, she has shown me how to be a philosopher.

I also thank my readers Jesse Prinz and Hagop Sarkissian for their thoughtful engagement with my project.

I thank my sister Petra for cheering me on and being proud of me.

I am grateful to my partner Michal for the countless hours he spent with me discussing metaethics, challenging my arguments, and forcing me to explain myself. He is my favorite person to talk philosophy with. His confidence in me and love during graduate school have meant the world to me.

I am deeply grateful to my Mom and Dad for encouraging me to pursue graduate school in philosophy. I thank them for raising me in a house bursting with books, ideas, and lively discussions. I thank them for the way they fostered my intellectual curiosity. Their unconditional love made this achievement possible.

For my beloved parents Richard and Bonnie Frank.

TABLE OF CONTENTS

INTRODUCTION	1
CHAPTER 1: The Queerness of (Naturalist) Moral Realism	12
1. Naturalist Moral Realism	12
1.1 Why Moral Realism?	16
2. Objections to Moral Realism	19
2.1 Queerness Objections to Moral Realism	22
2.2 Epistemological Queerness	25
2.3 Metaphysical Queerness: Supervenience	27
2.4 Metaphysical Queerness: Causal Efficacy and Explanatory Power	28
2.5 Metaphysical Queerness: The Normativity of Moral Facts	33
3. The Internalism Objection	37
4. The Authority Objection	42
5. Hume's Law and Deriving 'Ought from Is'	45
6. The Naturalistic Fallacy	51
7. Moral Disagreement and Relativism	60
8. Conclusions	67
CHAPTER 2: The Internalism Objection	70
1. Varieties of Internalism	72
2. The Case for Internalism	77
2.1 Phenomenology Argument for Internalism and Response	78
2.2 Oddness Argument for Internalism and Responses	82
2.3 The Connection Argument and Responses	87
2.3.1 There is No Reliable Connection	92
2.3.2 Alternatives to <i>De dicto</i> Desires	95

2.3.3 Defending <i>De dicto</i> Desires and Embracing Moral Fetishism	104
3. Two Kinds of Externalism	110
3.1 Desiderata for Externalism	115
3.2 Pluralistic Externalism	117
4. Conclusion	122
CHAPTER 3: Moral Indifference and Amoralism	123
1. A Brief Taxonomy of Moral Indifference	126
1.1 Amoralism	126
1.2 The Occasional Amoralist	128
1.3 Diversity in Moral Motivation	131
1.4 Depression and Listlessness	135
2. Internalist Responses to Indifference	137
2.1 Restricting the Scope of Internalism	140
2.2 Amoralists do not make moral judgments	143
2.2.1 Smith's Response to Amoralism	150
2.2.2 Prinz's Amoralist	153
2.3 The Impossibility of Community Amoralism	156
2.4 The Amoralist Is Motivated	160
3. Pluralistic Externalism and Moral Indifference	162
3.1 An Objection to Modal Intuitions and a Response	163
CHAPTER 4: The Authority and Normativity Objections	165
1. The Authority Objection	165
1.1 What has Authority?	168
1.2 What is Authority: Inescapability, Overridingness, or Just Normativity?	169
1.2.1 Inescapability	170
1.2.2 Overridingness	172
1.2.3 Normativity	175

1.3 Distinguishing Authority from Motivation	177
2. The Authority Objection: Korsgaard's Challenge	182
2.1 Responding to Korsgaard's Challenge	188
2.2 Is Korsgaard Conflating Justification and Motivation?	189
2.3 Prichard-Style Response to Korsgaard	192
2.3.1 Korsgaard's Response to Prichard	194
2.3.2 Evaluation of Korsgaard's Response to Prichard	196
2.4 Korsgaard Cannot Answer her Own Normative Question	198
2.4.1 First Person Requests	199
2.4.2 Why Should I Value My Humanity?	200
2.4.3 From Inescapability to Normativity?	201
3. Parfit's Normativity	203
3.1 Parfit's Objections to Non-analytical Naturalism	203
3.2 Parfit's Normativity Objection to Non-analytical Naturalism	206
3.3 Responding to Parfit's Normativity Objection	210
3.4 Parfit's Triviality Objection to Non-Analytic Naturalism	214
3.5 Responding to Parfit's Triviality Objection	217
4. Blackburn's Quasi-Realism	222
4.1 Why Quasi-Realism?	223
4.2 Quasi-realist Objectivity	224
4.3 The Problem with Quasi-realist Objectivity	228
5. Moral Realism and a Minimal Understanding of the Authority of Morality	232
CONCLUSION	237
REFERENCES	249

Introduction

Our days are filled with evaluative judgments. We might wake up and notice that the coffee is good today, think that the person who gave up his seat to the elderly woman on the bus is kind, notice that the neighbors grow beautiful roses, evaluate a student's paper as deserving a 'C+' but decide to give them a 'B-' anyway, or conclude that the President's proposed military intervention somewhere is justified. We might think about whether or not we should keep a date with a friend, even though we are tired and just want to go home and watch our favorite television show, where we sympathize with the fictional villain, while at the same time judging his behavior to be reprehensible. These judgments are sometimes kept to ourselves, but are often discussed with others, who may agree or disagree. Many of the countless evaluative judgments that we constantly make are also specifically moral judgments.

Moral realism is the view that moral judgments either succeed or fail in corresponding to a mind independent moral reality. The moral facts about people, actions, states of affairs, and so on and the moral properties of these people, actions, states of affairs, exist regardless of what we think of them or whether or not we are even aware of them. This vague characterization of moral realism immediately presents several puzzles. First, if we are committed to naturalism, whether methodological, epistemological, or metaphysical, it might seem puzzling how such moral facts and properties could be accommodated without embracing a realm of *sui generis* or mysterious moral facts and properties into our ontology and a corresponding faculty with which to perceive them.

The position I defend, naturalist moral realism, is a variety of Cornell realism. Cornell realism gets its name from the view's three most well know proponents: Richard Boyd, Nicholas Sturgeon, and David Brink. Boyd and Sturgeon were both professors at Cornell University and Boyd received his Ph.D. from Cornell. The version of Cornell realism that I defend claims that moral properties and facts exist objectively; this means that they are not dependent on or constituted by what we think about them. At the same time, these moral properties and facts are constituted by basic physical properties and facts.

This version of realism explicitly addresses worries about the incompatibility of moral facts and properties with natural facts and properties by characterizing moral facts and properties as being constituted by or supervening on basic physical facts and properties, in an analogous way to the relationship between the psychological, sociological, or historical facts and properties and basic physical facts and properties. On this view the psychological, sociological, or historical properties and facts can be understood as natural, without being characterized as physical properties and facts.

The relationship of supervenience is notoriously difficult to define. When one property A, supervenes on another property or set of properties B, there cannot be a change in A without a corresponding change in B. In the case of moral properties, the supervenience claim is just that something (an act, person, situation, etc.) cannot differ in its moral properties without simultaneously differing in its underlying physical properties. It is beyond the scope of this essay to discuss the supervenience relation in depth, for this reason, I will often characterize the relationship between the moral properties and facts and natural properties and facts as a relation of constitution.

Cornell realism is also the most promising candidate for a plausible version of realism because it suggests that there are non-analytic identities between moral properties and natural properties. These identities do not have to do with truths of meaning. Instead they are to be understood along the same lines as the property identities that we have discovered between things like water and H₂O.

At the same time, Cornell realism leaves open the possibility that the identities between moral properties and facts and natural properties and facts may be reductive or nonreductive. Nonreductive realism says that that moral properties and facts are constituted by natural properties and facts, rather than being reduced to natural properties and facts. Saying that the relationship between the moral and the natural is non-analytic is different from saying that it is nonreductive. A non-analytic version of naturalist moral realism does not involve a claim about equivalent meaning between moral terms and natural terms (or sets of terms). Nonreductive realism in contrast, rejects the claim that moral properties and facts are reducible to natural properties and facts. On such a view, the moral properties and facts can be both constituted by natural properties and facts but distinct from them in the way that the property of being a table is a distinct property from the properties of the microphysical constituents of the table. I favor nonreductive realism, but I have left that debate fairly untouched in this essay.

Even if realism can be understood in a way that is satisfactorily compatible with the commitments of naturalism, it still faces a host of objections. A set of persistent challenges to realism are articulated in J.L. Mackie's *Ethics Inventing Right and Wrong*, specifically in his section on the queerness objection to objective values. Mackie's set of queerness objections to the existence of objective values continue to be cited and

contemporary challenges to realism grounded in the queerness of moral facts and properties abound. I defend Cornell realism against two sets of persistent objections, objections based on moral motivation (the internalism objection) and objections based on the authority or normativity of morality.

In Chapter 1, I introduce Mackie's classic queerness objections to the existence of objective values. Mackie's central point, very roughly, is that if objective values were to exist, they would be a kind of thing like nothing else that exists. Part of the force of Mackie's argument comes from an assumption that moral properties and facts (objective values, in Mackie's language) could not be reconciled with a world of natural properties and facts. Cornell realism purports to be able to do just that. But even if it can, Mackie's queerness objections still pose a set of challenges to naturalistic moral realism in general.

Mackie's argument from queerness contains several sub-arguments, as he and others have noted. I present some of the central ways that queerness can be understood and concentrate on the strain of the argument having to do with metaphysical queerness. Metaphysical queerness can be further understood to have to do with the supervenience of the moral on the natural, the causal efficacy of objective values, or the normative force of objective values. I focus on the normative force strain of Mackie's argument. I further parse that objection into an objection having to do with the nature of moral motivation and an objection having to do with the authority of morality.

The objection based on the nature of moral motivation says that because realism is committed to cognitivism it has to construe moral judgments, either the sentences we

utter with moral content or the mental states we have with moral content, primarily as beliefs. The objection claims that problem with that is that beliefs cannot motivate us to act on them all by themselves. But this is inconsistent with the data, the objection claims, moral judgments *do* motivate us to act, all by themselves. This view is motivational judgment internalism (internalism for short). The objection concludes that realism is less likely to be the correct view because the cognitivism it is committed to is incompatible with internalism. The correct metaethical view, the argument continues, can allow for motivation to be built into moral judgments themselves, candidates include expressivism, projectivism, emotivism, and so on.

The objection based on the authority of morality says that realism cannot capture, or is incompatible with, the unique way that morality has authority over us. Morality is not the sort of thing one can opt out of; it binds us all regardless of our wishes to the contrary. This objection claims that realism somehow eliminates the “ought-ness” of morality, sometimes called its normativity. Christine Korsgaard’s formulation of this objection identifies realism itself as the problem, whereas Derek Parfit’s formulation of the objection identifies naturalism as the source of the problem.

I argue that although often conflated, objections to realism based on concerns about the nature of moral motivation are distinct from objections to realism based on the worry that it cannot capture the authority of morality. Both objections claim that there is something about realism (natural or non-natural) that is unable to easily accommodate or is inconsistent with an important belief that we have about the way morality works or an important fact about morality. One of the central projects of the essay is to distinguish the objection from motivation and the authority objection from each other,

understand them, explain how they are taken to pose a challenge to realism, and to defend naturalist realism.

Chapter 1 also considers two historically significant objections to moral naturalism: David Hume's alleged is-ought gap and G.E. Moore's naturalistic fallacy. I explain these arguments and why they are not a problem for Cornell realism. I address these objections because versions of them continue to appear in the contemporary metaethical literature on naturalism. For example, Parfit's objection to non-analytical naturalism raised in Chapter 4, mirrors the open question argument.

Chapter 1 ends with a brief discussion of a final argument that Mackie makes against objective values, the objection from disagreement. Mackie distinguishes it from the queerness objection, but the objection is often raised against realism in conjunction with the queerness objections, so it merits addressing.

Chapter 2 focuses on the objection to realism based on motivation, which I call the internalism objection. Internalism is the view that moral judgments necessarily, inherently, or essentially motivate the agents who make them to act on them. Whereas externalism is the view that moral judgments do not necessarily, inherently, or essentially motivate the agents who make them to act on them. Instead, the connection between making a moral judgment and having a moral motivation is a contingent one. First, I explain the case for internalism and some of the important distinctions between various types of internalism.

I then defend a pluralistic externalist account of moral motivation. On this view, at times, we are motivated to act on our moral judgments by sources like compassion and sympathy. We are also motivated to act on our moral judgments by conative states like

desires, which have content related to the moral judgment we make. For example, I might desire to do the right thing. Another person or at another time, someone may desire that their friend be happy. Another time a person may act out of a mostly or purely non-cognitive affective state like sympathy for the suffering person.

I argue that pluralistic externalism meets three important conditions that any theory of moral motivation should meet. First it should be consistent with the existing psychological and neuroscientific evidence about how the human motivational system works in general (the empirical evidence condition). Pluralistic externalism is an *a posteriori* theory and so is open to being amended based on scientific evidence, unlike many competing internalist theories, which posit an *a priori* necessary connection between moral judgment and motivation.

Second, a theory of moral motivation should be flexible enough to be compatible with the wide range of variability that we find in moral motivation between individuals and within a single individual across times and situations (the flexibility condition). Pluralistic externalism can embrace a wide range of types of motivation and can thus also easily explain a wide range of ways in which one can fail to be motivated as well. Pluralistic externalism is also the view that is most compatible with the phenomenon of moral indifference.

Third, it should be impartial with respect to ethical theories about what kind of motivation is morally praiseworthy; it should be as normatively neutral as possible (the neutrality condition). Normative neutrality is a traditional criterion for a successful metaethical theory. Pluralistic externalism is agnostic on which type of motivation represents the most morally, or the only morally praiseworthy type of motivation.

Further, it is neutral about whether or not motivation is something that can be morally assessed at all or whether it is rather the action that comes out of our deliberative process that truly matters morally. Although I do not embrace a strict division between metaethical and normative theories, I take neutrality on this issue to be important because I take the truth and nature of motivational externalism to be an empirical theory of moral psychology. So while how our motivational system actually works is relevant to what should be counted as morally praiseworthy or worthwhile, moral psychology should try to limit the extent to which it builds those kinds of views in.

Chapter 3 considers the phenomenon of moral indifference in detail, outlining some of the most important and widely discussed forms of moral indifference and amoralism. Those who are morally indifferent, like amoralists, are people who appear to be able to make moral judgments without being motivated to act on those moral judgments. If such people are conceivable then that is a serious counterexample to a *priori* internalism. If such people actually exist then the counter example touches a *posterior* forms of internalism as well.

I press the view that moral indifference is a far more commonplace phenomenon than may be currently recognized, especially when we consider the variability in levels of moral motivation between individuals and across an individual over time. I consider various internalist responses to moral indifference, especially the internalist strategy of reinterpreting what the morally indifferent person says to show that they do not really make a moral judgment at all. I show the internalist responses to be lacking and I also explain why pluralistic externalism can best make sense of moral indifference.

Chapter 4 addresses the authority objection to realism. The authority that morality has over us is very hard to pin down or define (for an important historical source on the authority of morality see Butler 1726). It is described variously as morality's normativity, the to-be-doneness of moral facts, the categoricity of morality, its overridingness, and its reason giving force, to name a few. The chapter first canvasses some of the ways the objection has been made and considering how we can understand what is meant by this mysterious power morality has over us. In surveying these various formulations, I argue that there is no reason to think that any one of them is necessarily incompatible with naturalist moral realism.

The authority objection to naturalist moral realism is motivated by several distinct metaethical positions. On some versions of the objection, the problem that the objection focuses on is that when thinking about moral properties as mind independent features of the world, we misplace the central connection they have to guiding our decision making. This is meant to be understood in something stronger than in a motivational sense. Korsgaard's version of the authority objection makes something like this claim; realism can never answer what she calls the "normativity question." But the authority objection also comes from nonnaturalist realists like Parfit. On this version of the objection, the problem of authority only arises when we consider versions of realism that also aim to be descriptive or natural.

Chapter 4 explains that, for Korsgaard, every metaethical theory must be able to explain, in a first person context, why someone should do what is morally required of them. Realism, she claims, cannot answer it sufficiently, while her Kantian constructivism can. I argue that Korsgaard's authority objection can be overcome and

respond to it in four related ways. Realism has stronger answers to the normative question than Korsgaard considers. At the same time, it seems that some of the force of the normative question may be stemming from a subtle conflation of justification and motivation. I also argue that Korsgaard's alternative to realism, Kantian constructivism, cannot adequately answer the challenge she poses. In fact, all metaethical theories have difficulty answering her normative question. This is an issue that I revisit at the end of the chapter. Realism, however, fares better than many others on the grounds that it captures objectivity in a unique way.

The second half of the chapter discusses Parfit's version of the authority objection. He thinks that morality will lose its normativity in the reason giving sense, if we construe moral facts as constituted by natural facts. Parfit's arguments against non-analytical naturalism, "the normativity objection," "the triviality objection," and "the fact-stating argument," all aim to show that natural facts by themselves cannot be normative. Instead, only a *sui generis* realm of nonnatural normative facts is consistent with normativity. Parfit's arguments against non-analytical naturalism do not succeed in showing that the normativity of reasons is lost in naturalist moral realism. The arguments that he takes to be strongest are only persuasive if one is already convinced that normativity must to be non-natural.

Having argued that neither paradigmatic cases of the authority objection to naturalist moral realism succeed, Chapter 4 concludes with a discussion of the kind of authority that naturalist moral realism does have. I introduce an account of minimal authority through a discussion of Simon Blackburn's quasi-realism. I draw a contrast between the kind of objectivity that moral realism can secure and the kind of pseudo-

objectivity that non-realist theories can secure. I chose Blackburn's view because it comes as close as an antirealist theory can come in successfully capturing many of the features of morality that we hope a metaethical theory can capture while remaining ontologically parsimonious.

The authority of morality is not the same as objectivity. But looking closely at one part of Blackburn's view, it becomes easier to see the kind of authority that realism can capture that non-realist views, like constructivism or quasi-realist expressivism cannot. I end the chapter by discussing what kind of authority is the best we can hope for, that is objectivity. This is something the realist can secure. This is a particular success for naturalist moral realism if at the same time it is able to do all of this without positing the existence of *sui generis* moral properties or facts.

Chapter 1: The Queerness of (Naturalist) Moral Realism

1. Naturalist Moral Realism

“Moral realism” describes a range of metaethical views. Most moral realists defend some or all of the following five claims:

1) the mind independence of morality: the claim that moral facts or properties exist objectively, regardless of what agents happen to believe about their existence or content;

2) ontological realism: the claim that moral facts or properties exist;

3) psychological cognitivism: the claim that moral judgments are beliefs;

4) semantic cognitivism: the claim that sentences that contain moral predicates can be true or false, that moral sentences express propositions, or that moral terms refer to moral properties; and

5) the rejection of error theory: the claim that sometimes our beliefs accurately represent the world and that sometimes our moral propositions are literally true.

Various formulations of moral realism emphasize different sets of these five conditions (Sayre-McCord 1986 p. 6; FitzPatrick 2009 p. 747). My emphasis is on defending the mind independence of morality, ontological realism, and psychological cognitivism.

There is no easy way to settle upon a definition of naturalism or of what makes a particular domain of inquiry a naturalistic one. One competing conception of naturalism which I will not rely on says that only those areas are natural which “are the proper object of natural scientific study” (Shafer-Landau and Cuneo 2006 p. 211). In the present context, naturalism can be understood, first, as an epistemological position that

rules out *a priori* knowledge about morality (Devitt 2010 p. 188). Second, moral naturalism is a metaphysical claim, that the “ethical facts and properties are exhaustively constituted by natural ones” (FitzPatrick 2009 p. 750). Another way of putting this is that everything that exists is natural. Together, moral realism and naturalism constitute naturalist moral realism.

The type of moral realism that I will defend is often referred to as Cornell Realism (Sturgeon 1985; Railton 1986; Boyd 1988; Brink 1989). This type of moral realism says that moral claims “purport to describe the moral properties of people, actions, and institutions,” and that some moral claims are true (Brink 1989 p. 7).

Cornell realism does not look for analytic identities or synonymy between moral terms and natural terms. Instead, it contends that although moral terms and natural terms “have...different meaning[s], these terms refer to the same (natural) propert[ies]” (Tännsjö 2010 p. 64). This view makes the primary focus of the metaethical theory metaphysical and psychological rather than semantic. The naturalist element of Cornell realism has to do with the relationship between moral properties and facts and natural properties and facts, not moral sentences or predicates and natural or descriptive sentences and predicates.¹

This view can be contrasted with analytical naturalism, for example, Frank Jackson’s analytical descriptivism, which says that “moral predicates and sentences could be replaced without significant loss by purely descriptive predicates and sentences” (Nuccetelli and Seay 2013 p. 133). Jackson describes the contrast; “we say, and the Cornell position denies, that, at the end of the day, we can say all there is to say

¹ While there is a difference between properties and facts, I use them interchangeably throughout and do not think that very much hangs metaphysically on whether what is real are moral properties or moral facts. I tend to think it is both, understanding facts as states of affairs.

about ethical nature in descriptive terms” (Jackson 2000 p. 146). I have chosen not to defend analytical naturalism, first, because it has the drawback of primarily focusing on moral language and grappling with a range of open question style arguments. Analytical naturalism and the debate surrounding it focuses extensively on mining linguistic intuitions which I do not think are particularly helpful in discovering what the correct picture of moral reality is. Our linguistic intuitions can only tell us about how we use moral words and form moral sentences; they can give us very limited insight into moral ontology or even moral psychology (Zimmerman 1980 p. 641). Second, I am skeptical about the notion of analyticity entirely and Cornell realism is not committed to defending the analytic/synthetic distinction in the way that analytical naturalists are.

Realism that is naturalist and non-analytic can come in either reductive or nonreductive forms. This distinction between reductive and nonreductive captures a metaphysical question about the relationship between the natural properties and facts and the moral properties and facts. There are Cornell realists on both sides of this divide.²

The kind of Cornell realism that I favor is nonreductive, in that moral properties and facts are constituted by, or supervene on natural properties and facts. Thus moral properties and facts are both irreducible in a sense, and natural at the same time. The distinction between reductive and nonreductive realism requires some unpacking. Brink characterizes his position as a nonreductive one. The important difference between a reductive position and a nonreductive position is that on a reductive position, moral properties and facts are identical to natural properties and facts. Whereas on the nonreductive view, moral properties and facts are constituted by natural properties and

² David Brink is a nonreductive naturalist. Richard Boyd is a reductive naturalist as is Peter Railton.

facts (Brink 1989 p. 9). Brink provides the following characterization of his nonreductive position:

Moral properties can be natural properties, though, even if they are not identical with natural properties. F can be G even if the property (or properties) designated by 'F' is not (or are not) the same as that (or those) designated by 'G'. If G actually composes or realizes F, but F can be, or could have been, realized differently, then G constitutes, but is not identical with F... a table is constituted by, but not identical with, a particular arrangement of microphysical particles, since the table could survive certain changes in its particles or their arrangement (Brink 1989 p. 157-158).

One reason for favoring nonreductive realism and the constitution relation rather than reductive realism and the identity relation is that nonreductive realism leaves room for the plausible claim that moral properties are multiply realizable (Brink 1989 p. 158).

That is, the same moral facts and properties could be constituted by more than one set of natural facts and properties.

Alternatively, a reductive realist would say that there is an identity between moral properties or facts and natural properties or facts that should be understood "on the model of other common scientific identity claims, such as water = H₂O, temperature = mean kinetic energy..." (Brink 1989 p. 157). The possibility of these kinds of identities between moral properties and natural properties has relied heavily on work on identity in metaphysics and philosophy of mind, especially the work of Saul Kripke and Hilary Putnam (Kripke 1980; Putnam 1992).

Even though I favor nonreductive naturalism, I address arguments that target both reductive naturalism and nonreductive naturalism, in part because there is so much variability in how this distinction is understood. For example, Russ Shafer-Landau self identifies as a nonnaturalist realist, but is characterized by many others as a nonreductive naturalist (Bedke 2012 p. 111fn). At the same time, someone like Brink

who presents his view as nonreductive says that moral properties are “nothing over and above” natural properties and so is sometimes characterized as a reductive naturalist (Bedke 2012 p. 111fn). It seems that if one is a nonreductive naturalist, nearly all of the same objections apply, or are posed against one’s view, as when one is a reductive naturalist, and *visa versa*.

The realist position that I defend is often accompanied by the claim that moral properties and facts can play a causal and explanatory role. “[M]oral properties make a genuine and ineliminable contribution to the best explanation of experience (in particular, of moral belief)” and “moral truths [...] pull their weight in naturalistic explanations”(Hooker 2008 p. 587). The causal explanatory role of moral properties is sometimes used as evidence that moral properties exist and used as evidence that these properties are natural properties, since if they were not natural or constituted by natural properties it would be very hard to understand how they could interact with natural properties and cause things to happen in the natural world. (See sections 2.3 and 2.4).

1.1. Why Moral Realism?

There is strong *prima facie* evidence to support moral realism; it is the position most consistent with several elements of our everyday moral thinking and language (Smith 1995; Blackburn 1984; Brink 1989; Mackie 1977; Shafer-Landau 2003). For example, cognitivism, one important element of moral realism, captures the way we think about our moral judgments. When we assess whether an action is right or wrong, a person’s character is virtuous or vicious, or a state of affairs is good or bad, we take

ourselves to be aiming at representing the way things really are, the moral facts or properties of the action, person, or state of affairs. We also tend to think that we can engage in genuine moral disagreements with other people. If I claim that 'eating meat is morally wrong' and you claim that 'eating meat is morally permissible,' we take ourselves to be asserting conflicting claims and we assume that we cannot both be correct (See Ross 1939 p. 30-41).

Nor do we think that if everyone in a society, or even everyone in the world, came to believe that deliberate cruelty to children is morally permissible that it would actually become morally permissible. There is no dependence relation between what we think about morality and what morality actually requires of us. This is not to say that if we were psychologically constituted in a radically different way morality would not be different. For example, if we weren't the sort of beings that could feel pain, then it may no longer be the case that pinching another person for fun would be wrong, holding everything else equal.

We also tend to think that it is possible that an individual agent or a large group of agents could be in moral error about one of their moral judgments or about a whole host of them. Entire societies can be swept up in morally abhorrent ideologies. Even if every single person thinks they are right, they are still wrong.

Finally, we believe that it is possible to make moral progress. For example, we tend to believe that a country which has overcome racial segregation is morally better than the state it was in when segregation was standard.

These are just examples, meant to suggest the way that realism is embedded in ordinary moral thinking and practice. Thus, consistent with other moral realists, I claim that the burden of proof falls on those who would deny moral realism.

Debate about burden of proof, discussed most commonly in epistemology and legal scholarship, focuses on what it is constituted by, whether it exists at all, and how to determine who it falls on in an argument. The burden of proof is often assumed to be on the position that denies a claim that counts by itself as common sense or is consistent with common sense in the way the opposing position is not.

A possible preliminary objection to realism at this point, is that our everyday moral thought, language, and practice do not support moral realism, thus attempting to undermine the claim that antirealism carries the burden of proof. One form of this challenge is that folk morality is not truly realist, instead it has a cultural relativist streak in it (Sarkissian et al. 2011; Björnsson 2012). Sarkissian et al. acknowledge that several studies suggest that people do generally take morality to be objective and that they understand moral disagreements to be about something other than the interlocutors opinions or feelings (Nichols 2004; Goodwin and Darley 2008). However, Sarkissian et al. argue that recent studies that frame the questions respondents are asked in a different way, they produce different results. They claim that when participants consider the moral judgments of people who are radically different from them (either in terms of their culture or their “values or ways of life”) “their intuitions move steadily toward a kind of relativism” (Sarkissian et al. 2011 p. 486). That is, they tend to judge that it is possible for their own moral judgments to be correct and at the same time for the person who

holds conflicting moral judgments to be correct, or they hold that there is no objective fact of the matter at all in such cases.

Ideally, realists should have ways to respond to these results and be able to defend their contention that folk morality supports objectivism and thus puts the burden of proof on those who would deny it, but this debate is not the focus of this essay. One plausible avenue for a realist response is to scrutinize the way that the experiment questions that generated the relativist responses were framed. For example, people have a strong impulse not to want to seem judgmental or culturally imperialist. At the same time, they may have been taking the descriptions of alternative value systems which are built into the questions as clues that in order to “get the answer right” they should take note of these facts and respond accordingly. Another second possible objection to the idea that antirealism carries the burden of proof in the debate is to argue that the ethical appearances of everyday moral discourse and thought should not carry the weight that realists give to them (Singer 2005; Björnsson 2012). There are problems with this dismissal however (Sandberg and Juth 2011).

2. Objections to Moral Realism

Realism faces several pressing objections that suggest that despite appearances to the contrary, there are no moral facts or properties in the world at all. In the past twenty five years, objections to realism and alternative theories, like Simon Blackburn’s quasi-realism, Gilbert Harman’s relativism, Allan Gibbard’s expressivism, or Christine Korsgaard’s Kantian constructivism, have become increasingly sophisticated, and hold the promise of being able to explain the moral appearances and give us a satisfying

sense of objectivity without any of the metaphysical baggage of realism (Gibbard 1992; Blackburn 1993; Harman and Thomson 1996; Korsgaard 1996).

But Cornell realism remains the most persuasive version of realism and more persuasive than the closely related antirealist metaethical alternatives, like quasi-realism, error-theory, and constructivism, because it can avoid two major sets, very broadly construed, of objections to which realism has been traditionally vulnerable.

The first set of objections comes from the charge that objective values (moral facts and properties) are “queer” (“queerness objections”). This objection can be understood in many different ways. Several of these readings will be discussed in the following sections. But the strain of the objection I will centrally focus on in this essay is the metaphysical queerness of the normativity of objective values. I argue that this objection to moral realism can be broken down into two distinct sub-types, the motivation objection and the authority objection. These objections will be discussed in section 2.5.

The second set of objections are directed at naturalism specifically, these arguments are rooted in G.E. Moore’s naturalistic fallacy and the purported “is-ought” gap. These objections will be discussed in sections 5 and 6.

The first set of objections that moral realism needs to contend with is J.L. Mackie’s (and other’s) queerness objections. These are objections that claim that if objective moral values did exist in a mind independent way, they would be unacceptably different from anything else that exists (Mackie 1977). Many contemporary challenges to moral realism are a species of Mackie’s classic argument from queerness.

Part of Mackie's set of objections to moral realism come from the charge that moral realism is inconsistent with a naturalist picture of the world. Cornell realism is most capable of diffusing this objection by espousing naturalism compatible with realism. Naturalist moral realism can avoid the charge of being inconsistent with science by arguing that moral facts and properties are realized or constituted by physical facts and properties. Moral facts and properties merely operate at a different level of description than the basic sciences, in much the same way the special sciences, history, economics, and sociology do. While moral facts must depend on physical facts, they need not reduce to physical facts (Brink 1989; Sturgeon 1985; Boyd 1988). So why should moral properties and facts be thought to be queer if they are really just natural properties and facts, rather than *sui generis* or mysterious properties and facts?

The objection cannot be diffused that easily. Queerness objections capture a mystery about the normativity of objective values (moral properties or facts) that realism must address. Not only do queerness objections challenge the compatibility of objective values with a naturalist worldview, they also challenge the compatibility of objective moral values with the normative force of morality. How to understand this objection will be discussed in section 2.5.

At the same time, espousing naturalism is often thought to leave realists open to a second set of objections, the charge that they violate Hume's law, attempting to bridge the gap between is and ought, or commit Moore's naturalistic fallacy (see sections 5 and 6). Contemporary naturalist moral realists respond to these objections by explaining that they do not intend to provide analytic identifications of moral terms with natural terms or *a priori* reductions of moral language to non-moral language. Instead, they

suggest that moral facts and properties are constituted by or supervene on more basic physical facts and properties. This makes moral facts and properties no less real, and no less natural, than the facts of history, psychology, or biology (Brink 1989). Beyond this chapter I will not directly address the naturalistic fallacy or the is-ought gap again. I address them in this chapter because of their historical importance for naturalist moral realism and because of their ubiquity in the metaethical literature.

I argue that none of these objections succeed in disproving naturalist moral realism, or even shifting the burden of proof from antirealism, non-cognitivism, or relativism, to realism. I will start with Mackie's queerness objections. Then discuss two other important challenges to realism, the naturalistic fallacy and arguments based on an is-ought gap. I end the chapter with a brief digression on arguments from moral disagreement to relativism.

2.1 Queerness Objections to Moral Realism

Objections to moral realism often involve disputes about the existence of moral properties, facts, beliefs, or judgments. Ever since Mackie's argument for error theory and against objective values, this broad class of objections is usually referred to as arguments from "queerness" (Mackie 1977 p. 38).³ Mackie, however, was by no means the first to raise these objections against objective values or moral realism. These arguments can be found before him in Moritz Schlick, C.L. Stevenson, and A.J. Ayer (Schlick 1939; Stevenson 1937; Ayer 1936/2012).

³ Not everyone who has made these objections to realism was influenced by Mackie by any means. But his oft cited and vague objection contains the seeds of a long standing, complex, and difficult set of problems at the intersection of moral philosophy, metaphysics, and moral psychology.

Queerness objections depend on undermining the likelihood of the existence of objective values (moral properties or facts). Usually, that is done by highlighting just how strange moral properties, facts, beliefs, or judgments would have to be when contrasted with other types of properties, facts, beliefs, or judgments. Critics of moral realism argue that such moral entities, if they existed in the sense that realists says they do, would be very strange things and this is taken as a reason to doubt that they exist at all (Finlay 2007 p. 6). Mackie's argument contains several different sub-arguments, not all of which are fully fleshed out. In fact, there may be nothing distinctive about the argument from queerness at all (Shepski 2008 p. 374-375). Instead, it may just be a vague and colorful way of repackaging several old objections to realism having to do with authority, normativity, and motivation. Mark Platts writes about Mackie's argument; "[t]he queerest thing about this as it stands is the claim that it is an argument...The world is a queer place (Platts 1980 p. 72).⁴

What is Mackie actually accusing the existence of objective values of when he calls them "queer" or "of a very strange sort, utterly different from anything else in the universe"(Mackie 1977 p. 38)? First it could be that queerness simply means unnecessary in terms of explaining our moral practices. This is similar to Harman's arguments that our behavior could be explained just as well without positing the existence of any moral facts and instead just citing the beliefs that we have about morality and our dispositions and emotional reactions, etc. (Shepski 2008 p. 375; Harman 1977). See section 2.4.

Second, it could be that by "queer," Mackie means "nonsensical" (Shepski 2008 p. 375-376). This interpretation is doubtful, however, since Mackie rejects this point

⁴ See also Judith Jarvis Thomson "The Legacy of Principia" (Thomson 2003).

when considering what Hare says about moral language that aims at “express[ing] a statement of fact, and yet we have no idea what would be the case if it were true, the sentence then that sentence is (to us) meaningless” (Hare 1952 Sec. 1.4). Embracing such an argument would be inconsistent with the way that Mackie sets up his queerness objection. Mackie “is asking us to conceive the entities in question and to reflect on just how queer they would be if they existed” (Shepski 2008 p. 376). So surely he does not think that objective values are like round squares in this way, that is, inconceivable (Shepski 2008 p. 376).

A related interpretation of queerness is that Mackie thinks objective values, were they to exist, would be incomprehensible (Shepski 2008, p. 376). This is a weaker claim than saying that objective values are nonsensical. Instead the claim is that they are beyond the grasp of human understanding. But as Lee Shepski points out, this is hardly a good reason to reject the existence of something. Surely there could be things that exist that our beyond the ken of our limited minds (Shepski 2008 p. 376).

A fourth way to understand queerness may be that the objective values are mysterious in the sense of being unexplainable. Of course, not being able to currently explain something, such as is the case in physics, cannot rule out its existence or make it “in principle inexplicable” (Shepski 2008 p. 376).

Fifth, and finally, perhaps queerness means that objective values would have to be *sui generis*, unique. That is, they either have or lack some property that everything that exists has or they have some property that nothing else that exists has. This is the most likely explanation for what Mackie means by queerness. But this claim will also require some unpacking. In what sense does Mackie take these values to be unique

(Shepski 2008 p. 377)? The property that objective values would have or either lack could be one that has to do with the way in which we have knowledge about them, their epistemic status, or what they fundamentally are, their metaphysical status.

Mackie famously claims:

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. 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. 95-6).

In the first sentence, the claim being made is that moral facts or properties are metaphysically different from everything else in the natural world. Whereas, the second sentence makes an epistemological point, that given the characteristics he assumes that moral facts or properties would have to have, they are not the sorts of things we could find out about in the ordinary empirical ways. Instead, we would have to posit the existence of a special way of getting access to them. Thus at base, the two strains of the argument are metaphysical and epistemological (Sinnott-Armstrong 2007; Shepski 2008; Sinnott-Armstrong 2010).

2.2 Epistemological Queerness

Mackie argues that if objective values existed our only way of knowing about them would be through a special faculty of intuition. Unfortunately, we have no evidence that we have this sort of faculty.

When we ask the awkward question, how can we be aware of this authoritative prescriptivity, of the truth of these distinctively ethical premises or of the cogency of this distinctively ethical pattern of reasoning, none of our ordinary accounts of sensory perception or introspection or the framing and confirming of explanatory hypothesis or inference or logical construction or conceptual analysis, or any combination of these will provide a satisfactory answer; 'a special kind of

intuition; is the lame answer...to which the clear headed objectivist is compelled to resort (Mackie 1977 p. 39).

So it seems that the criteria that Mackie is establishing here for an entity to be able to escape a charge of queerness is that it can be investigated empirically. But, according to Mackie, "objective values cannot be the sort of things that are accessible to empirical investigation" (Shepski 2008 p. 379).

The first kind of response that will be obvious for a naturalist to make here is that what Mackie calls objective values, what might be more easily called moral facts or properties, are constituted by or identical to natural facts and properties. Because of this there is no need to posit a special faculty of intuition with which to discover them. Instead, objective values are discoverable through the reputable means of empirical investigation or perhaps some kind of Rawlsian reflective equilibrium (Schmidt 2012 p. 124-125; Brink 1989; Sturgeon 1985; Devitt 2002). Nor does the use of intuitions in ethics commit one to the existence of a mysterious faculty of intuition. For example,

...reflective equilibrium is the method of the sciences as well as ethics. In both cases, use of this method in critically evaluating our beliefs is seen as inevitably involving some appeal both to the evidence of experience and to currently held substantive theory—for example, application of a principle of inference to the best explanation requires recourse to our going theory as well as experience in assessing the plausibility of competing explanatory claims. Thus, the fact that moral epistemology cannot dispense with an appeal to existing moral judgments or "intuitions," and can subject such judgments to criticism and revision only in broadly coherentist ways, does not show a fundamental difference or discontinuity between moral and scientific epistemology (Darwall, Gibbard, and Railton 1992 p. 169).

The use of intuitions seems only to become mysterious and queer in philosophy if we assume that intuitions are giving us access to the pure *a priori* realm. This is not the case, if instead, we think about intuitions as a kind of *a posteriori*, "empirically based expertise" (Devitt 1996 p. 76).

Mackie anticipates the second kind of response that the realist might give and calls it the “companions in guilt” strategy (Mackie 1977 p. 39). The strategy is to point out that there are other kinds of things that we do not think can be discovered empirically, like the findings of mathematics, but this does not make us doubt their existence (Shepski 2008 p. 380; Sinnott-Armstrong 2007 p. 46). Mackie’s reply to companions in guilt is skepticism that there are truly other things that can be discovered non-empirically. If we think there are entities that are truly not amenable to empirical explanation “then they too should be included, along with objective values, among the targets if the arguments from queerness” (Mackie 1977 p. 39).

The third response to Mackie’s charge of epistemic queerness is that even if one grants his claim that the only way to gain knowledge of moral facts is through a faculty of moral intuition, which does not exist, this does not establish the conclusion that objective values do not exist. “The only way to get from ‘no knowledge’ to ‘no fact’ is via a premise that we would be able to know moral facts if there were any. But Mackie gives us no reason to accept such a premise”(Sinnott-Armstrong 2007 p. 46). There may be moral facts to which we have no access (Enoch 2009).

2.3 Metaphysical Queerness: Supervenience

Another way that Mackie characterizes the strange features of objective values is by questioning the kind of relationship that they could have with natural facts and properties. If one is a naturalist realist, then again, the reply here comes easily. Moral properties are constituted by natural properties and so in this way are no different from what exists in the rest of the world. This response will run into a related metaphysical

worry that Mackie has about supervenience and that several other later opponents of realism have taken up (Blackburn 1988, 1971).⁵ What kind of relationship would natural facts have to have with moral facts? Mackie asks, for example,

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... 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. 41)?

As several philosophers have pointed out, Mackie asks these questions and seems to assume that there will be no ready answers (Sinnott-Armstrong 2007 p. 50). Mackie’s questions here do not aim to straightforwardly defend the claim that the relationship between natural facts and moral facts must be “queer.” Instead he is merely attempting to shift the burden of proof onto naturalist realists who suggest that this is what the relationship is like. So responses to this objection would have to come in the form of substantive explanations of the nature of the constitutive or supervenient relationship between natural and moral facts or properties (Brink 1984; Shafer-Landau 1994; Railton 1986).

2.4 Metaphysical Queerness: Causal Efficacy and Explanatory Power

A related kind of metaphysical queerness has to do with the causal role that moral properties play in the natural world. This is an objection that Mackie hints at when he asks, of “anything that is supposed to have some objective moral quality, how this is linked with its natural features?” (Mackie 1977 p. 41) Harman has taken up this

⁵ Blackburn’s version of this argument against naturalist realism depends in part on a revamped version of the open question argument and in part on a principle he calls the “ban on mixed worlds” (Blackburn 1988).

argument, defending the position that moral facts and properties, do not play a causal role in our forming the moral beliefs or judgments that we come to have and make (Harman 1977, 1977, 1986).⁶ Nor do moral facts and properties explain anything that happens in the natural world. Harman asks, what do moral facts add to any explanation of an event that cannot be captured by our moral beliefs?

Although there is a superficial similarity between the way that we evaluate scientific theories and the way we evaluate moral theories, Harman claims that moral theories, unlike scientific theories, cannot be observationally tested. In the scientific case, we test our theories against observations, while in the moral case we test our theories against our judgments about specific cases. However, this is not the same kind of testing. We must suppose facts about protons and other entities postulated by science, in order to best explain our observations of phenomena in the world.

Harman's widely discussed example asks us to imagine a physicist in a lab, who sees a vapor trail and makes the judgment "there goes a proton" (Harman 1986, 1977). In the contrasting ethical case, someone walks down an alley and sees some kids lighting a cat on fire and makes the judgment "that's wrong." Harman claims that in order to explain the physicist's judgment we need to cite the existence of the proton, that is, to suppose that there really was a proton. Reference merely to the physicist's psychology, beliefs, etc. will not explain her observation. In contrast, the judgment in the cat burning case, of "that's wrong," can be explained solely by reference to the observer's psychology or moral sensibility. We need not suppose that the burning of the cat is actually wrong because the wrongness does not need to be cited in the explanation of the observation.

⁶ This debate has also recently been taken up by Brian Leiter (Leiter 2001).

It would seem that all we need assume is that you have certain more or less well articulated moral principles that are reflected in the judgments you make, based on your moral sensibility. It seems to be completely irrelevant to our explanation whether your intuitive immediate judgment is true or false (Harman 1977 p. 7).

In the physics case, the scientific theory that the physicist holds, which predicts that we will observe a proton under certain circumstances, explains the physicist's observation "there goes a proton." The observation counts as evidence in favor of whatever scientific theory best explains that observation, but this is not the case in morality.

The observation of an event can provide observational evidence for or against a scientific theory in the sense that the truth of that observation can be relevant to a reasonable explanation of why that observation was made. A moral observation does not seem, in the same sense, to be observational evidence for or against any moral theory, since the truth or falsity of the moral observation seems to be completely irrelevant to any reasonable explanation of why that observation was made (Harman 1977 p. 7).

Nicholas Sturgeon has several forceful realist responses to Harman's claim that moral facts or properties lack a causal and explanatory role (Sturgeon 1985).⁷

Essentially, Sturgeon argues that Harman is begging the question against the realist in at least two ways.

First, Cornell realists argue that moral facts do play an obviously explanatory role in the best explanation of moral judgments and other non-moral facts.⁸ For example, Hitler's depravity, his immoral character, helps explain the Holocaust. In part, it was Hitler's moral depravity that caused the deaths of six million Jews (Sturgeon 1985).

⁷ A different way of responding to Harman's challenge, that Sturgeon does not pursue, would be to argue that although moral facts do not play an explanatory role, we still have reason to believe that they exist. This is David Enoch's strategy in "An Outline for an Argument for Robust Metanormative Realism" (Enoch 2007). Shafer-Landau also makes the argument that being able to play a causal explanatory role is not the test for whether something exists or not (Shafer-Landau 2003).

⁸ Many naturalists also think that the fact that moral properties play a causal role in the natural world gives us a good reason to think that they are also natural properties. When we make an identity between heat and molecular motion we are in large part relying on the identity of causal roles. Even if we cannot make a reductive identity between moral and natural properties, locating the causal roles of moral properties shows us something about the kinds of properties that they are.

Ethical properties often play a causal role in our lives; we are regularly helped by other people's good character and harmed by their bad characters. Revolutions and regime change can be explained in part by the injustices of the government or the system itself (Railton 1986 p. 190-193).

Brink points out that sometimes injustice, bad character, etc., can have a causal effect before the agents that are being harmed are even aware of it, and before they have formed moral beliefs that their system of government is economically unjust, with great inequalities between the rich and poor, for example. Even if a person believes that their society is just, they may begin to feel "unreflective resentment about [their] own social positions" and "begin to sympathize with other disadvantaged members of [their] class" (Brink 1989 p. 189). Eventually these feelings lead the person to reflect, form moral judgments, and take political action. Cornell realists argue that unless you make antirealist assumptions to begin with, these will sound like plausible explanations.

Second, Harman does not deny the naturalist moral realists claim that moral properties supervene on or are constituted by natural properties, in Harman's example, "[t]he badness of the cat burning supervenes on other natural properties" (Sturgeon 1985). Harman claims that the fact that we judge the action of setting the cat on fire to be wrong can be explained without reference to any moral properties or facts, all we need to reference are our own psychologies.

But if the wrongness supervenes on natural properties, or is constituted by those natural properties, then if the property of wrongness would not have been there, then the physical properties of the situation would have to be different as well. If the physical properties were different, for example, setting a cat on fire did not cause the cat pain or

death, we probably would not make the judgment that it was wrong. In other words, if the kids' act had not been wrong it would have lacked the features that made it wrong such as being an act of deliberate, intense and pointless cruelty. But, if as Harman contends, the moral wrongness of the situation is completely irrelevant to the explanation of the judgment, then the judgment would have obtained even if there had been no fact of wrongness. It is strange for Harman to claim that the fact that the children's act was wrong is completely irrelevant to our judgment that it was wrong.

Sturgeon asks the question, would we have still made the judgment "that's wrong," if the act hadn't actually been wrong? Sturgeon makes the assumption that having the property of being a deliberate, intense, pointless act of cruelty *makes* an act wrong. So, relying on this assumption Sturgeon asks, if the act had not been a deliberate act of pointless cruelty, would we have still thought it was wrong? Harman argues that we do not need to suppose the existence of moral facts in our explanations of our moral beliefs. Natural (non-moral) properties alone are enough to explain our judgments. Harman is asking 'if the act had been exactly the same in all of its natural properties (and all other facts had been the same) and yet the moral facts had been different (the act had not been morally wrong) would we still have thought that their act was wrong?' But for a naturalist moral realist, it is hard to even imagine the moral facts and all of the other facts coming apart in the way that Harman supposes they can. While Sturgeon recognizes that Harman would not allow for his interpretation of the question, his interpretation remains plausible.

If we allow Harman's interpretation of the question, then he is faced with a third and final objection. Harman's skeptical argument can apply just as well to other

theoretical entities (such as protons), the external world, or other minds, as to moral facts and properties. There is no reason to see Harman's argument as posing a special problem for moral facts or properties and their explanatory or causal roles. To be able to imagine that an act was an act of intense, pointless, cruelty, and yet was not morally wrong, has to involve accepting that our whole moral theory is completely wrong.

Harman admits that, "[o]bservations are 'theory laden.' What you perceive depends to some extent on what theory you hold, consciously or unconsciously" (Harman 1977 p. 4). But Harman excludes the possibility of relying on any background moral theory when we think about whether or not we would have still thought the children's act was wrong (Loeb 1998). To be consistent, Harman cannot allow any background scientific theory to enter into the question "would the physicist still have thought there was a proton there even if there was not?" If we cannot rely on background microphysical theory then we must suppose that there could have been a vapor trail in the cloud chamber and yet no proton. That means that the physicist, relying on her observations, could have made the judgment that there is a proton there, even if there was no proton there.

Third, the moral realist can argue, as Shafer-Landau does that moral facts and properties do not need to meet Harman's standard of having causal power in order to be counted as real facts and properties (Shafer-Landau and Cuneo 2006 p. 224-225).

2.5 Metaphysical Queerness: The Normativity of Moral facts

A preliminary type of the queerness of objective values I have considered is queerness understood as explanatory redundancy, as Harman has argued. The next

section will focus on the queerness of objective values understood in terms of their normative force.

One of the queer features that moral properties, facts, judgments, and beliefs are said to have is their normative force; broadly speaking this is their action-guiding nature.⁹

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*. An objective good would be sought by anyone who was acquainted with it, not because of any contingent fact that this person, or every person, is so constituted that he desires this end, but just because the end has to-be-pursuedness somehow built into it. Similarly, if there were objective principles of right and wrong, any wrong (possible) course of action would have not-to-be-doneness somehow built into it. Or we should have something like Clarke's necessary relations of fitness between situations and actions, so that a situation would have a demand for such-and such an action somehow built into it (Mackie 1977 p. 40) [italics added].

Nothing besides moral beliefs, judgments, properties, or facts seems to have this kind of normative force. The objection to moral realism based on the claim that moral properties, facts, beliefs, or judgments have normative force can also be interpreted in several ways (Sinnott-Armstrong 2010 p. 56-57).

So far as ethics is concerned, my thesis that there are no objective values is specifically the denial that any such categorically imperative element is objectively valid. The objective values which I am denying would be action directing absolutely, not contingently (in the way indicated) upon the agent's desires and inclinations (Mackie 1977 p. 29).

The quote above characterizes objective values as having the "categorical imperative element" and as being "action directing absolutely." Categorical imperatives are rules that we must follow regardless of our contingent desires or goals. They command

⁹ Other kinds of judgments such as, economic, mechanical, & prudential judgments may also seem to have normative force, depending on how normative force is interpreted. Here I will restrict normative force to moral judgments.

absolutely and apply to us because we have the capacity to reason. Categorical imperatives do not have to do with an agent's motivation for acting. This suggests that an important target of the normative force part of the queerness objection is the authoritativeness of objective values. In other words, Mackie is puzzled by the existence of an "objectively binding ought" (Bedke 2010 p. 43).

On the other hand, when Mackie writes about action-guidingness as having to do with an agent's desires and inclinations this seems to have to do with the connection between objective values and our motivations to act on those objective values. The above cited passage shows that Mackie has both of these elements in mind, although he does not explicitly distinguish between them.

As mentioned above, one way to understand the normative force objection is in terms of motivation and the other is in terms of the authority of morality.¹⁰ Later Mackie makes a claim that targets naturalism specifically, including a comment on both the motivational power of morality and the authority of morality.

On a naturalist's analysis, moral judgments can be practical, but their practicality is wholly relative to desires or possible satisfactions of the person or persons whose actions are to be guided; but moral judgments seem to say more than this. This view leaves out the categorical quality of moral requirements...both naturalist and non-cognitive analysis leave out the apparent authority of ethics (Mackie 1977 p. 33).

On the motivation focused view, the normative force of moral judgments, for example, motivates us to perform the acts we take to be moral. Critics of moral realism charge that the realist picture of morality is inconsistent with the way moral judgments seem related to moral action. Thus Mackie is critiquing what anyone proposing the existence of objective values must be suggesting; that something exists which

¹⁰ Sinnott Armstrong takes the two relevant interpretations to be between "reasons internalism" and "motivational internalism" (Sinnott-Armstrong 2010 p. 57).

necessarily motivates and that is also mind independent. This strain in the argument can be seen when Mackie writes: "...values differ from natural objects precisely, in their power, when known, automatically to influence the will" (Mackie 1977 p. 40).

The second way of understanding the normative force objection has to do with the authority of morality.¹¹ This latter objection I will call the authority objection.¹² On this view, the normative force of moral judgments or properties explains why we should be moral or how objective values provide authoritative reasons for action. The authority of morality can be understood as the categorical nature of moral demands, the importance that moral demands have in directing our lives, or the way in which moral considerations override all others, among other interpretations. On Mackie's view, although our ordinary ways of thinking and speaking about morality assume that morality is both objective and prescriptive in the authoritative sense, no such thing could exist.¹³ Of course, Mackie's position would be stronger here if he presented an independent justification for thinking that nothing objectively prescriptive could exist (Shepski 2008 p. 378).

Critics of moral realism argue that realist accounts cannot capture the way that we experience moral demands as demands on action. In part this is because critics see

¹¹ A similar distinction is also made in David Brink's discussion of Kant's categorical imperative between "inescapability, authority and supremacy" (Brink 1997). Finlay makes a similar distinction between "motivational queerness" and "normative queerness" (Finlay 2007 p. 15). Also see Richard Joyce's discussion of Mackie's queerness argument for a similar distinction (Joyce 2007).

¹² One of the earliest sources on the authority of morality is Joseph Butler, (*Fifteen Sermons*, Sermon II, III "Upon the Natural Supremacy of Conscience" 1726/1749).

¹³ Here it seems like Mackie must be assuming that only internal reasons exist, rather than external reasons as well. This is not a point that I think that realists should necessarily use to rebut Mackie, but it is still important to recognize. He is assuming that to be a reason for me to act it has to in some way help me realize my subjective set of desires. On the opposite view, reasons can be of this sort, or they can be of another sort, there can be categorical reasons for action, for example, that are completely independent of one's subjective desires. I don't think that naturalist moral realists have to rely on defending the idea of external reasons to defend the position that values can be both objective and prescriptive in the sense of being authoritative.

moral facts or properties (especially if they are natural facts or properties) as just sitting there in the world, unable to connect in the right way with human reason and action. And so it is always possible to wonder why such facts or properties are relevant to us or our lives, why should we care about them or pay attention to them at all (Korsgaard 2003). In the same way that it is optional for us to pay attention to or care about other kinds of facts (the gaseous composition of the sun, for example).

The motivation objection and the authority objection are two distinct objections. Moral motivation has to do with psychological questions about how moral facts, properties, and judgments influence our actions. This is distinct from the authority of morality, which has to do with the metaphysical status of morality. The difference between these two objections is similar to a distinction that is sometimes made between motivating reasons and normative reasons.¹⁴ Motivating reasons have to do with an agent's goals or desires, they psychologically explain an agent's actions (Smith 1995 p. 96). Whereas normative reasons have to do with moral requirements and justification (Smith 1995 p. 95).

Out of the normative force strain of the queerness objection, I will focus on these two objections to objective values, the motivation objection (the internalism objection) and the authority objection.

3. The Internalism Objection

The objection from motivation starts by pointing out that most versions of moral realism involve a commitment to psychological cognitivism. This is the view that moral judgments express beliefs and that moral thoughts purport to represent the world.

¹⁴ There are also explanatory reasons which help us understand why people do what they do.

Opponents of moral realism often argue against moral realism by arguing that cognitivism is false. If cognitivism is false, then realism has to be false, too.

In order to show that cognitivism is false, opponents of realism argue that it is incompatible with the combination of two other claims about how moral judgment and motivation interact. The first of these claims is the view that in order for someone to have a motivation to act, they must have a desire or attitude, not just a belief. For example, just *believing* that the object in front of me is a cookie will not get me to eat the cookie. I must also *desire* or *want* to eat the cookie.

The second claim in the argument against cognitivism is that when we make a moral judgment it necessarily or inherently gives us a motivation to act on it. When I *judge* that sending money to Oxfam is the right thing to do, I am necessarily *motivated* to send money to Oxfam. So, moral judgments cannot be beliefs, because moral judgments motivate us to act. Therefore, cognitivism is false, and realism, which is assumed to be inseparable from cognitivism, is also false.¹⁵

Someone hoping to defend naturalist moral realism has three ways to proceed in responding to this objection based on motivation. To defend naturalist moral realism she can reject (a) cognitivism, the view that moral judgments express beliefs and moral thoughts purport to represent the world, reject (b) the view that moral judgments necessarily or inherently motivate, also called internalism, or reject (c) the view that in order for someone to have a motivation to act they must have both a belief and an attitude or desire, also called the Humean theory of motivation (Brink 1997 p. 6).

¹⁵ Rowland Strout argues that this type of objection is a kind of “extension of Moore’s open question argument”. If one takes Hume’s purported claim that “motivation is internal to judgment” then “you cannot truly judge something to be right or good without thereby having some motivation to act or feel accordingly. If that is right there should be no room for the following version of an open question: ‘I can see that this is good, but why should I care?’” (Strout 2010 p. 854).

The moral realist may reject (a) and instead argue for a version of realism that includes a different theory of moral thought and language, either a non-cognitivist or hybrid view. This route has not been popular, but a growing number of versions of realist emotivism and realist expressivism have recently been defended.¹⁶

The moral realist can also reject (c), the view that motivation requires a belief/desire pair. To make a case for this, the realist must argue that when properly understood, some beliefs *do* provide motivation in themselves. The rejection of (c) is an option pursued by rationalists, but not often by *naturalist* moral realists.¹⁷

Finally, the moral realist may reject (b) the view that moral judgments necessarily or inherently motivate. That is, they can reject internalism. In this way it is possible for the moral realist to capture both the admittedly central role that moral judgments play in our decision-making and the appearance of the motivational pull of moral judgments. This is the view I will defend.

The first part of my strategy in defending naturalist moral realism is to undermine the internalism objection. In order to clarify, the internalism objection (the argument from motivation), I introduce some terminology. The opponents of realism argue that the following three positions are inconsistent: (a) Cognitivism, (b) Internalism, and (c) The Humean theory of motivation (Miller 2003 p. 219). I will briefly explain each one of these three positions in order to show why opponents of realism claim that there is an inconsistency between the three positions.¹⁸

¹⁶ See e.g. (Brown 2008; Copp 1995).

¹⁷ See e.g. (Nagel 1970; McDowell 1988; McNaughton 1988; Dancy 1993; Wiggins 1990). An exception is Smith (1994) who is a rationalist, realist, and a naturalist.

¹⁸ Michael Smith calls the conflict between these three positions *The Moral Problem* (Smith 1994). Smith's own solution however, is to reinterpret internalism and the Humean theory of motivation so as to make the set of three positions consistent.

(a) Cognitivism is a view about both moral psychology and about moral language. As a view about moral psychology, cognitivism is the position that moral judgments are primarily beliefs and not conative states, such as pro-attitudes or desires. As a view about language, cognitivism is the position that moral language is truth apt. Here, truth aptness can be characterized in various ways, as, for example, that moral sentences express propositions that can be evaluated as true or false or that moral terms refer to moral properties.

Cognitivism is usually defined in contrast to non-cognitivism. Non-cognitivism is the view that moral judgments are conative states such as desires, emotions, or pro-attitudes, and not beliefs. On semantic non-cognitivism, moral language is not truth apt. Moral language expresses conative states, which do not assert that something is the case. According to semantic non-cognitivism, moral terms, such as 'good' and 'right' do not refer to moral properties.

(b) Internalism, in its most general form, is the view that moral judgment is necessarily, essentially, or inherently tied to motivation.¹⁹ Types of internalism and arguments for them will be distinguished in Chapter 2. For now, the kind of internalism that I will focus on is motivational judgment internalism. Motivational judgment internalism is the view that moral judgments by themselves, (that is, without the help of other mental states), necessarily, essentially, or inherently provide the agent making the judgment with a motive to act on them. This view is contrasted with motivational judgment externalism, the view that moral judgments are only contingently connected

¹⁹ W.D. Falk is credited with first making the distinction between internalism and externalism (Falk 1947). The roots of this distinction can also be found in Harold Arthur Prichard in *Duty and Interest* (1928).

with motivation to act morally and that moral judgments are not necessarily, essentially, or inherently motivational.

(c) The Humean theory of motivation says that beliefs by themselves are motivationally inert; desires are required to move people to act.²⁰ This theory is widely accepted among moral philosophers, including most naturalist moral realists, and often taken to be a piece of common sense. The view is that beliefs alone are incapable of getting an agent to act. Only a belief that is hooked up the appropriate way with a desire can move an agent to action. As applied to ethics, it follows that the belief that 'giving to charity is the right thing to do', cannot in itself cause me to give to charity. I must also have a desire; one obvious candidate for the corresponding desire is the desire to do the right thing. Sometimes the Humean distinction between beliefs and desires is explained in terms of the metaphor of "direction of fit." Whereas beliefs are supposed to "fit" with the world, the role of a desire is to get the world to "fit" with it.

The distinction between directions of fit is originally made in John Searle's discussion of intentionality and can be applied to both speech acts and mental states. Beliefs have a mind-to-world direction of fit (Searle 1979/1985). This means that the intentional content of the mental state "is supposed to fit, match or correspond to, etc. the world" (Fotion 2000 p. 40). For example, the belief that "The sky is blue" has a mind-to-world direction of fit because it is a state of the world, that is, the sky being blue that makes that belief true. The belief aims to fit with the way the world is. Desires, hopes, wishes, etc. have a world-to-mind direction of fit, in that these mental states aim to change the world in a way that fits with the mental state. When the world changes to fit with a desire, for example, that desire is satisfied. We can also understand this by

²⁰ Whether or not this is a good characterization of Hume's own view is not at issue here.

saying that beliefs and wishes or desires have different “conditions of satisfaction;” “[b]eliefs are satisfied when what is believed is true, intentions satisfied when what is intended is done, wishes satisfied when they are fulfilled” (Fotion 2000 p. 102).

The objection to moral realism based on moral motivation exploits the fact that these three positions, cognitivism, internalism, and the Humean theory of motivation, seem to be mutually exclusive. This is because, the cognitivist attests, moral judgments are beliefs. Beliefs, by themselves, cannot motivate an agent to act. However, there is an inherent connection between making a moral judgment and having a motivation to act. This must mean that moral judgments cannot be beliefs.

The objections to naturalist moral realism based on the role of moral motivation all share questionable assumptions, so I will treat them as roughly one class of objections. And from now on, when I use the term “internalism objection,” I will thereby refer to all of the relevant objections to naturalist moral realism based on motivational reading of normative force. I will defend moral realism in two main ways. I do this first by attacking some of the arguments for or considerations that are supposed to weigh in favor of internalism. Second, I defend it by arguing for the alternative view, externalism, which is in short, the view that moral judgments need not intrinsically or necessarily motivate an agent to act on them (See Chapters 2 and 3).

4. The Authority Objection

According to the authority objection, moral realism wants to have its cake and eat it too. The objection is used against moral realism generally and also used specifically against naturalist moral realism. According to moral realism, obligations and

duties exist, regardless of the beliefs people have about them. The authority objection questions the authority that any set of facts or properties, natural or not, could have over our behavior. In other words, the moral realist seems to lack a convincing response to the moral skeptic who asks, 'why should I be moral?'²¹ The puzzling upshot of this objection is that, while the realist may be able to secure the objectivity of moral rules and the truth of moral facts, they cannot secure its authority.

The authority objection comes up in several specific forms against naturalist moral realism. This is because, *naturalist* versions of moral realism insist that the objective moral facts and properties that they contend exist, do not introduce anything mysterious that we should worry about into our ontology, nor do they commit us to anything epistemologically troubling, like a special faculty of intuition with which to detect them. Instead, moral properties and facts are just like historical, sociological, or psychological facts. We have access to them in the ordinary way, through observation. And they are constituted by or supervene on more basic physical facts and properties.

But, the objector protests, morality is not like history or psychology. The facts of history and psychology are not facts that have any authority over us. They don't have to matter to us in the way morality does. They are not action guiding or practical. I can shrug and walk away from them in a way I cannot shrug and walk away from moral facts and properties. In Parfit's version of the authority objection against naturalism the charge is that when we understand moral facts as natural facts we lose their normativity entirely; normativity drops out of our picture of the world. While many facts that have normative relevance remain, for example, facts about human suffering and happiness,

²¹ One of the earliest sources on the authority of morality is Joseph Butler (*Fifteen Sermons*, Sermon II, III "Upon the Natural Supremacy of Conscience");(Butler 1726).

what disappears is the *sui generis* fact that something is what I have most reason to do. Surely the realist cannot deny this prescriptive element in moral thought and discourse. The idea of authority is still vague at this point; it might mean the inescapability of moral requirements, the normativity of morality, the special level of importance that morality has in directing our decision making and life choices, the way in which moral considerations seem to override all others, or the categorical nature of moral reasons.²² Without narrowing down authority at all, the following is the basic argument from authority:

- (1) Morality has a certain kind of authority over us.
- (2) (Naturalist) Moral realism either precludes this kind of authority or moral realism cannot adequately explain this kind of authority.

Therefore, (naturalist) moral realism must be false.

Two obvious strategies emerge for responding to this argument. First, the realist can deny premise one, arguing that, in fact, morality does not have the kind of authority that critics claim it has. Second, the realist can give an account of the authority that comports with naturalist moral realism.

My strategy will employ both of these responses. Premise one is true, given a certain, limited account of authority. The first step here is to consider what critic might mean by the authority of morality. I consider two plausible accounts of morality's authority. The first type of authority I consider is authority as "normative force," as given in Korsgaard's rationalist neo-Kantian constructivism. I present her normativity based objection to moral realism and respond to it.

²² Sometimes the authority of morality is also explained in terms of the notion that "moral requirements provide reasons for acting which override other kinds of reasons"(McNaughton 1988 p. 115).

The second objection I consider focuses specifically on *naturalist* moral realism. I examine Parfit's objection to naturalism which says that when we understand moral properties and facts as natural properties and facts they lose their ultimate reason giving force.

Third, I examine Simon Blackburn's quasi-realism, which is a position that claims that projectivism combined with a minimalist theory of truth can capture everything we want from a realist metaethics, without the metaphysical baggage that realism brings with it. I argue that without the metaphysical baggage, Blackburn's theory cannot secure the authoritativeness of morality.

Contrary to prominent accounts of authority, I argue for a minimal authority for morality based in objectivity. Given the minimal account of authority, premise two is false, this minimal kind of authority is no more a problem for naturalist moral realism as a metaethical theory than any other metaethical theory.

Before continuing in the next chapters to focus on the internalism objection and the authority objection, it is important to canvass four other objections to naturalist moral realism. There are broader ways to interpret the argument from queerness or to argue against moral realism, naturalist or not, that have not yet been discussed. In the following three sections I will discuss objections to realism based on the alleged is-ought gap, Moore's naturalistic fallacy, and moral disagreement.

5. Hume's Law and Deriving 'Ought from Is'

The gap between is and ought that Hume is said to have pointed out is used as evidence in a wide range of arguments for non-cognitivism and antirealism. It is used to

support the claim that there is an insurmountable gap between descriptive and evaluative domains, thus specifically targeting naturalist versions of moral realism.

However, as we will see, the existence of this gap does not provide evidence against realism whether it is of a naturalist or nonnaturalist stripe. Hume famously wrote:

In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought*, or *ought not*, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it ... (Hume 1739-40/2000 Bk. 3, Pt. 1, Sec. 1).

This brief passage contains what many refer to as the fact/value gap, the no ought from an is rule, or Hume's Law. R.M. Hare is largely responsible for giving this passage in Hume a central place in modern and contemporary metaethics (Hare 1954). A great deal of work has gone into interpreting this brief passage.

The simplest interpretation of Hume's passage is that he is making the logical point that a conclusion containing an ought (an evaluative premise) cannot be validly derived from premises that do not contain an ought (evaluative premise), premises that are merely descriptive (Frankena 1939 p. 467).²³

It is useful to employ a distinction made by Charles Pigden on the difference between at least two different interpretations of Hume's famous passage, one, that the passage supports the "logical autonomy of ethics" and two, that the passage supports the "semantic autonomy of ethics" (Pigden 1989 p. 128).

²³ Descriptive and evaluative language comes from (Searle 1964).

The logical autonomy of ethics means that “moral conclusions cannot be derived from non-moral premises”(Pigden 1989 p. 128). This is the simple logical point about the nature of deductive logic; “you cannot get out more than you put in” in order for an argument to be valid (Hattiangadi 2007 p. 40). It is important to note that, if Hume’s Law shows that no more than this, that moral conclusions can be deduced from entirely non-moral premises, this does not rule out the possibility of other forms of inference, “some nondeductive mode of inference, perhaps inductively” (Timmons 1998 p. 761).

But the logical point is not the only way that the passage has been understood. The semantic autonomy of ethics is the claim that there is “...a realm of *sui generis* non-natural predicates which do not mean the same as any natural counterparts” (Pigden 1989 p. 128). In other words, semantic autonomy is the idea that “you cannot decide moral conclusions of any kind (such as conclusions concerning virtue or featuring the right and the good) from entirely non-moral premises...” even “with the aid of analytic bridge principles” (Pigden 2010 p. 7). Analytic bridge principles are “truths of meaning” (Pigden 2013 p. 2). However, for nonreductive naturalist moral realists, the bridge principles in question would not have to be truths of meaning, instead they would be synthetic identities between moral facts and natural facts (Timmons 1998 p. 761).

The logical autonomy of ethics is a weaker position than the semantic autonomy of ethics. While the semantic autonomy implies the logical autonomy, logical autonomy does not imply semantic autonomy (Pigden 1989 p. 128). In response, cognitivists argue that non-cognitivism is unnecessary to explain the logical autonomy of ethics. Many kinds of talk are autonomous in this same way. As Pigden points out, we cannot conclude anything about hedgehogs from premises that do not include hedgehogs

(Pigden 1989 p. 132). Ethics is logically autonomous in same sense. When understood as a point about logic, Hume's law "does not imply that there is any fundamental difference between the moral and the nonmoral" (Pigden 2013 p. 3).

Arthur Prior challenged even this more trivial logical point, giving several influential counter-examples to the logical autonomy of ethics (Prior 1960). If the logical autonomy of ethics is false, then the semantic autonomy of ethics must be false as well.

- (1) Tea-drinking is common in England.
- (2) Therefore either tea-drinking is common in England or all New Zealanders ought to be shot (Prior 1960 p. 210).²⁴

Prior's argument shows that ethical conclusions *can* be deduced from premises with no ethical content. The response to this argument has been to point out the triviality of the types of ethical conclusions that can be drawn. Prior admits that there is something odd about all of these examples, but after seeing them (and modified versions of them), one can no longer deny that it is impossible to validly deduce conclusions about morality ("ought" conclusions) from non-moral ("is") premises alone.

However, non-cognitivists like Hare, take Hume's point to suggest the stronger position, the semantic autonomy of ethics that moral claims differ fundamentally from other types of claims (Pigden 2013 p. 3). Hare writes:

*No imperative conclusion can be validly drawn from a set of premises which does not contain at least one imperative....*In this logical rule, again, is to be found the basis of Hume's celebrated observation on the impossibility of deducing an

²⁴ Prior's other examples include:

- (1) Anyone who does what is not common in England ought to be shot.
- (2) All New Zealanders drink tea;
- (2) Therefore either tea-drinking is common in England or all New Zealanders ought to be shot.(Prior 1960 p. 201-2) .

As well as: (1) There is no man over 20 feet high

- (2) Therefore there is no man over 20 feet high who is allowed to sit in an ordinary chair (Prior 1960 p. 202).

'ought'-proposition from a series of 'is'-propositions -- an observation which, as he rightly says, 'would subvert all the vulgar systems of morality', and not only those which had already appeared in his day (Hare 1952 p. 28).

For Hare, for example, moral judgments are not descriptive, instead they are commands (Hare 1952).

John Searle directly criticized the semantic autonomy of ethics, focusing on the case of promise making (Searle 1964).²⁵ His example is as follows:

- (1) Jones uttered the words "I hereby promise to pay you, Smith, five dollars."
- (2) Jones promised to pay Smith five dollars.
- (3) Jones placed himself under (undertook) an obligation to pay Smith five dollars.
- (4) Jones is under an obligation to pay Smith five dollars.
- (5) Jones ought to pay Smith five dollars. (Searle 1964 p. 44).

In this counter-example, Searle is not attempting to show, strictly speaking, that a non-evaluative premise can entail an evaluative conclusion, but that there is "more than a contingent relation" between the non-evaluative and the evaluative (Searle 1964 p. 44). He focuses on a distinction between two different kinds of "objective facts," introducing "institutional facts" (Searle p. 55). Statements of fact like, 'this apple is red,' or 'the earth rotates around the sun,' are distinct from statements of fact that "contain[] words such as married,' 'promise,' 'home run,' and 'five dollars.' The latter, are facts that only exist as fact due to the existence of sets of institutions (Searle 1964 p. 54-5). Thus these are "institutional facts," which contrast with "brute facts" (Searle 1964 p. 55).²⁶

...a man gets married or makes a promise only within the institutions of marriage and promising. Without them, all he does is utter words or make gestures (Searle 1964 p. 55).

²⁵ Hare challenges Searle's analysis in *The Promising Game* (Hare 1964).

²⁶ So the brute fact would be that a man has a green piece of paper, but the institutional fact is that he has \$50.

The institutions of promise making, marriage, baseball, and money, are all constituted by a set of rules. These are rules that, “do not merely regulate but create or define new forms of behavior... they, as it were, create the possibility of or define that activity” (Searle 1964 p. 55). Promising is an “institutionalized form of obligation” regulated by a particular set of rules. While it is not a “brute fact” that “Jones ought to pay Smith \$5” it is an institutional fact. And this is how Searle proposes to violate Hume’s law and get from ‘is’ premise to an ‘ought’ conclusion.

I started with a brute fact, that a man uttered certain words, and then invoked the institution in such a way as to generate institutional facts by which we arrived at the institutional fact that the man ought to pay another man five dollars. The whole proof rests on an appeal to the constitutive rule that to make a promise is to undertake an obligation (Searle 1964 p. 56).

Whether or not Searle is correct that ethics is not semantically autonomous on these grounds, there is no independent reason to think that the ethics is semantically autonomous.²⁷

For the purposes of defending moral realism, it is important to see why non-cognitivists think that Hume’s Law supports their position (Hare 1952; Stevenson 1937; Nowell-Smith 1952; Flew 1986). There is no straightforward argument from Hume’s Law to the truth of non-cognitivism. Instead, as Francis Snare has argued, non-cognitivists actually have to assume the truth of their theory in order to espouse Hume’s law at all (Snare 2002).

Pigden suggests that non-cognitivists’ best strategy is a kind of inference to the best explanation, which he calls the “Generic Argument.” (Pigden 2010). It runs as follows:

²⁷ Hare and others have challenges Searle’s arguments and conception of institutional facts (Hare 1964).

- 1) You can't derive an Ought from an Is, moral conclusions from non-moral premises.
 - 2) The best explanation of 1) is
 - (a) that logic is conservative: in a valid inference you can't get out a new relation or affirmation that you haven't put in;
 - and
 - (b) that since moral judgments are fundamentally different from non-moral propositions [insert non-cognitivist theory of choice] moral words represent a 'new relation or affirmation'
- Therefore, probably:
- 3) Moral judgments are fundamentally different from non-moral propositions [insert non-cognitivist theory of choice] (Pigden 2010).

But as Pigden points out, the conservativeness of logic alone, premise (a), can explain the No-Ought-from-Is principle, and thus the non-cognitivist inference to the best explanation fails (Pigden 2010). This is because premise (b) is unnecessary. And so the conclusion that moral judgments are fundamentally different is unsupported.

It is also important to notice the genealogical connection between Hume's Law and the internalism objection to moral realism. In Hume, one of the important distinctions between moral judgments and other kinds of judgments is the connection that moral judgments have to the emotions and to action; "[m]orals excite passions, and produce or prevent actions" (Hume 1739-40/2000; Hattiangadi 2007 p. 41). It is, of course, a matter of historical debate whether or not Hume himself was an internalist about motivation (Brown 1988; Darwall 1992; Mele 1989).

Many philosophers wrongly assume that Hume's Law, is the same thing as Moore's naturalistic fallacy.²⁸ The next section will briefly present and dismiss the naturalistic fallacy as a serious threat to naturalist moral realism.

6. The Naturalistic Fallacy

²⁸ For an explanation of why they are not the same see (Frankena 1939).

A distinct argument against naturalist realism is Moore's naturalistic fallacy and open question argument. The naturalistic fallacy is committed when any definition of good is given that is non-normative. "The naturalistic fallacy occurs only when one confuses good...with any natural object whatever" (Moore 1903/1993 p. 14). It is not only naturalistic definitions of good that the naturalistic fallacy targets, it also targets what Moore called "metaphysical" definitions of good, for example, a definition offered by divine command theory that says that good is whatever God commands, Moore would say has committed the same fallacy as a definition of good that identifies goodness with being most evolved (Moore 1903/1993 p. 58). So the "fallacy is the same if one confuses any object with any other which it is not" (Baumrin 1968 p. 85). Moore writes:

Yet a mistake of this simple kind has commonly been made about good. It may be true that all things which are good are *also* something else, just as it is true that all things which are yellow produce a certain kind of vibration in the light. And it is a fact, that Ethics aims at discovering what are those other properties belonging to all things which are good. But far too many philosophers have thought that when they named those other properties they were actually defining good; that these properties, in fact, were simply not other, but absolutely and entirely the same with goodness. This view I propose to call the naturalistic fallacy and of it I shall now endeavor to dispose (Moore 1903/1993 Chpt. 10, Par. 3).

Although Moore was a realist, he was a nonnaturalist realist. He proposed that good is simple and indefinable, just like the color yellow is. Moore argues for the simplicity and indefinability of good by a process of elimination. If good is not simple and indefinable then it must be "either ...a complex, a given whole, about the correct analysis of which there may be disagreement; or else it means nothing at all" (Moore 1903/1993, Chpt. 1, Par. 13). The open question argument is then meant to rule out these other two options, that good is complex or that good is meaningless (Baumrin 1968 p. 86).

The hypothesis that disagreement about the meaning of good is disagreement with regard to the correct analysis of a given whole, may be most plainly seen to be incorrect by consideration of the fact that, whatever definition be offered, it may always be asked with significance, of the complex so defined, whether it is itself good...[one of] such proposed definitions, it may be easily thought, at first sight, that to be good may mean to be that which we desire to desire, Thus if we apply this definition to a particular instance and say 'When we think A is good, we are thinking that A is one of the things which we desire to desire,' our proposition may seem quite plausible. But if we carry the investigation further and ask ourselves 'Is it good to desire to desire A?' it is apparent, on a little reflection that this question is itself as intelligible, as the original question "Is A good? (Moore 1903/1993 Chpt. 1, Par. 13)

Moore concludes: 'what we desire to desire' cannot be the definition of good. This same pattern will inevitably hold for every complex that is suggested to provide a definition of good, whether it is utilitarian, based on individual emotions, evolutionary ethics, cultural norms or agreements, or the will and commands of a deity (Strout 2010 p. 855).

The naturalistic fallacy and the open question argument have been used against a wide range of views that Moore did not necessarily have in mind when he took this position originally (Baumrin 1968 p. 88). Naturalists must respond to these objections for at least two reasons. First, the most obvious reason is that they are open to being accused of having committed the naturalistic fallacy. Second, many non-cognitivists and other antirealists accept the naturalistic fallacy and use it as part of the evidence that their alternative theories better capture the nature of ethical judgments and language (Hare 1952; Stevenson 1937; Ayer 1936/2012). For example, early emotivists often accepted Moore's analysis of attempt to find definitions of "good." However, they rejected the notion that this must mean that good is *sui generis* and accessible only through the faculty of moral intuition. They offered their alternative analysis of how we

use moral language to express emotion or attitudes and to prescribe and command behavior rather than to describe the moral part of the world.

There are at least four major responses to the open question argument, first that it does not capture the correct criterion for what makes something a good definition of something else, second, that it begs the question against naturalism, third, is the no interesting analysis objection, and fourth, it confuses sense and reference, leaving room for naturalist moral realism to offer a promising account of identities or constitution relations between natural facts and properties and moral facts and properties.

The first immediate problems with Moore's open question test for any definition of good is that it is far too strong of a criterion for deciding whether a definition is accurate or not.²⁹ First we have no reason to think that having "doubts" about a definition being the correct one tells us that the definition is in fact incorrect (Baumrin 1968 p. 86). If this were the appropriate test for definitions, then people's doubts about "[m]any of the definitions of logic, mathematics, physics, psychoanalysis, etc., are dubious to the plain man (and even to those not so plain), but they are not on that account rendered inadequate" (Baumrin 1968 p. 86). When a definition of good is presented, for example, 'good is what we desire to desire,' it may always be coherent to ask, whether what we desire to desire is actually good. But just because the definition fails the open question test, we cannot on these grounds alone reject the definition.

The proper conclusion may be that we are quite confused...we may have 'two different notions before our minds' because we don't understand one or both halves of the definition, or because we misuse one or both halves, or because

²⁹ Baumrin offers this helpful reconstruction of the first two premises of the open question argument:
(1) Of every definiens in every definition it can be significantly asked whether what it refers to is or is not the thing, or has the property that we think is referred to by the definiendum.
(2) If we doubt that it is or does then the definiens does not render (for us) what the definiendum does (render to us) (Baumrin 1968).

we do not know how to separate the connotative overtones attaching to every word in the ordinary vocabulary of every natural language from their denotative application...but none of these, at any rate, show that we aren't now staring at the correct definition of good! (Baumrin 1968 p. 87)

The second general objection to the open question argument is that it begs the question against the naturalist realist (Frankena 1939). Frankena's discussion of the naturalistic fallacy is particularly helpful.

The charge of committing the naturalistic fallacy can be made, if at all, only as a conclusion from the discussion and not as an instrument of deciding it (Frankena 1939 p. 465).

A better way to understand the naturalistic fallacy is to understand the "generic fallacy which underlies it," which he calls the definist fallacy (Frankena 1939 p. 471). Frankena says the "definist fallacy is the process of confusing or identifying two properties, of defining one property by another, or of substituting one property for another" (Frankena 1939 p. 471). The fallacy is not specific to non-natural, natural, or ethical properties. The fallacy is committed anytime someone gives an account of two distinct properties as if they were one and the same property (Frankena 1939, p. 471). Strictly speaking, it does not really make sense to call the definist fallacy a fallacy either; rather it is simply a "mistake" (Frankena 1939, p. 472). If the definist fallacy is a better way to understand what the naturalistic fallacy is really about, then the question becomes: "do those who define ethical notions in non-ethical terms make this mistake" (Frankena 1939, p. 472)?

Those who attempt to give analytic definitions of good will contend that there are not two properties involved at all, so they aren't making this mistake. Instead, "what they are saying is that two words or sets of words stand for or mean one and the same property" (Frankena 1939, p. 472). The definist fallacy is the mistake of treating two

different properties as if they were the same property. Frankena argues that while Moore may be right that the property of “goodness is not identifiable with any ‘other’ characteristic (if it is a characteristic at all)...the question is: which characteristics are other than goodness, which names stand for characteristics other than goodness?” (Frankena 1939, p. 472). The problem is that Moore prejudices the question of whether or not good is indefinable and names any attempt to do so a fallacy (Frankena 1939, p. 473).

Specifically, Moore assumes that any identification of good with a natural property will fall to the open question test. These definitions or identities will always intuitively feel open. Further if the naturalist has found the correct definition then to be able to continue to ask, ‘yes X is N, but is it good?’ would reveal a “conceptual confusion” (Miller 2003 p. 15.). So the openness of the question is based on a preexisting assumption about nonnaturalism being correct and not offering an independent argument.

A related objection asks why so much weight should be given to the imaginary intuitions of Moore’s competent users. Moore or his defenders can claim that the doubts, or the openness of the questions Moore refers to, are not meant to give “definitive proof” that the proposed definition is not the correct one (Strandberg 2004 p. 182). Instead, Moore’s defenders might argue that “competent language users” intuitions about definitions should carry some weight:

If some truths hold merely in virtue of language, and if the existence of such truths is an essential feature of us being able to communicate, it seems reasonable to assume that competent language users possess knowledge of these truths ... it seems reasonable to treat their responses of doubt to questions such as "Is whatever is M also G?" as indications of "AT" not having the same meaning as "G" (Strandberg 2004 p. 184).

Thus on this response, the open question test is a legitimate test of meaning because there is a purported “tight connection...between analyticity and linguistic competence, and between linguistic competence and responses to the questions utilised in the argument” (Strandberg 2004 p. 184). But, *prima facie* evidence against a naturalistic definition was not what Moore was seeking, after all, he accused those of making the naturalistic identifications of committing a fallacy, not making a less plausible inference to the best explanation.

This brings us to the third, and related, major objection, which has been called the “no interesting analysis objection” and by others the “paradox of analysis” objection (Miller 2003 p. 16; Langford 1968). Moore’s open question test for any definition of good supposes that it is a necessary condition of a definition being correct that the question to be closed. But, paradoxically, if the question is closed, then the definition will be uninformative, uninteresting, and/or trivial.

Moore’s conditions of analytic adequacy require all analyses to be either trivial or false. An analysis states that an identity obtains between the object denoted by the analysandum and the object denoted by the analysans. If the identity does obtain then by the terms of Moore’s argument, the analysis is *trivial*, stating no more than a tautology. But if the analysis is nontrivial, ‘significant,’ there is no identity between the two and the analysis must be false (Hutchinson 2001 p. 31).

There are many counter-examples to this way of understanding what a correct definition or conceptual analysis is like, from math, logic, and philosophy (Miller 2003 p. 16).

Smith points out that once we realize that [t]here “are unobvious and informative” analysis of concepts, the openness of Moorean open questions looks less significant. Perhaps some questions are closed by our concepts, but are nonetheless cognitively open to us (Smith 1995 p. 39).

One way of describing the way that Moore understands 'good' is that he recognizes this paradox and thinks that in the case of good it cannot be resolved (Umphrey 2002 p. 34). Soames suggests that Moore may have assumed the principle of "The Transparency of Meaning"

If two expressions α and β , mean the same thing (e.g. express the same property), and in addition, an individual x (fully) understands both α and β then (i) x will know that α and β mean the same thing, and (ii), x will know that any two sentences (of the sort Moore is considering) that differ only in the substitution of one of these expressions for the other mean the same thing, and thereby express the same proposition (in the case of declarative sentences or question (in the case of interrogatives) (Soames 2003 p. 46-7).

There are counter-examples to this principle however.³⁰ Smith makes the point that this can be illustrated through the distinction between knowledge how and knowledge that (Smith 1995, p. 37-9). While someone might possess a skill, like being able to speak English or drive a car, this is not the same thing as being able to state what the rules of these activities are.

The fourth major objection is that Moore ignored the distinction made by Gottlob Frege between "sense and reference" (Frege 1892/1948 p. 17; Miller 2003 p. 17). The reference of a term is "the object for which it stands, either by having been assigned to that object or by uniquely describing it" (Mendelsohn 2005 p. 33). The reference of a term is something that we can discover empirically. Whereas the sense of a term, is "that 'wherein the mode of presentation is contained' and it thus carries the burden of introducing, presenting, or picking out a referent" (Mendelsohn 2005; Frege 1892/1948). It is possible to understand the sense of a term without knowing its referent, for example, the sense of the tallest building in Manhattan is clear to me, even though I am

³⁰ For more on this, see Smith 1995 (Smith 1995 p. 37-39). To find the distinction between knowledge how and knowledge that more generally, See: (Ryle 2009; Dummett 1996).

not aware that the building is lower Manhattan's Freedom Tower. In Frege's example, the terms Morning Star and Evening star had different senses, because people believed they were different stars. However, in fact had the same reference, they were both Venus.

Moore assumed that the open question test was a good test for the meaning of terms, based on their synonymy (Pigden 1991). However, synonymy is not a good guide to telling whether two terms refer to the same property because their sense and reference can come apart. So on this objection, what is going on with the open question test is say that someone proposes a utilitarian analysis of the concept good- that good is whatever maximizes happiness for all. We can coherently ask, but is it good to maximize happiness for all? Because good and maximizing happiness for all may have the same reference (if this was the correct identity) but not the same sense. In other words, Moore assumes that any statement identifying two properties can be correct if the two terms for the properties are synonyms.

This response is especially important for naturalist moral realism; it opens the door for naturalists to argue that the identities between moral properties and natural properties are like the empirically discoverable identities between two natural properties, like the identity between water and H₂O (Pigden 1991). This is the strategy taken by the Cornell Realists. For example, Brink writes:

The naturalist can concede that there are neither synonymies nor meaning implications between moral and non-moral, for instance, natural terms and still maintain that moral facts and properties are identical with or constituted by, natural and social scientific facts and properties. The naturalists' identity or constitution claims can be construed as expressing synthetic moral necessities (Brink 1989 p. 166).

Although there have been contemporary attempts to revitalize the open question argument they often they fall victim to one of these same flaws. For neo-open question arguments see: (Horgan and Timmons 1992, 1992). Otherwise they are guilty of simply begging the question against the naturalist by mining for intuitions about the openness of questions that are so abstract that they make the value of these intuitions highly questionable.

7. Moral Disagreement and Relativism

Although Mackie does not include disagreement in the section in his book on the queerness argument, it is an argument that Mackie makes against the objectivity of values and can be construed as a kind of queerness. It would be 'queer' if morals were objective, and yet such persistent disagreement about values between individuals, societies, and cultures remained.

Arguments based in disagreement specifically challenge the objectivity that realists' theories claim morality has. These arguments are often inferences to the best explanation to relativism, in Mackie's case to error theory. The argument is as follows. The phenomenon of widespread moral disagreement is something that a metaethical theory should be able to explain. The best way of explaining it is that ethics is not objective and that there are no ethical facts or properties for people to converge on. If such properties or facts existed objectively, we would expect less intractable and persistent disagreement than we currently have. Instead, disagreement can be best explained with a skeptical metaethical view, such as relativism, quasi-realism, subjectivism, non-cognitivism, or error theory. The argument is often made through a

contrast in the level of convergence that exists in ethics with the level that exists in sciences (taken to be the paradigmatic cases of objective areas of inquiry).

...the actual variations in moral codes are more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values (Mackie 1977 p. 37).

According to Gilbert Harman's metaethical relativism "moral truth and justifiability, if there are any such things, are in some way relative to factors that are culturally and historically contingent...it is about the relativity of moral truth and justifiability" (Wong 1991 p. 442). He proposes relativism as the best explanation for persistent ethical disagreement (Harman and Thomson 1996; Harman 1975).

The main reason to believe there is not a single true morality is that there are major differences in the moralities that people accept and these differences do not seem to rest on actual differences in situation or disagreements about the facts. It is hard for me to believe that all moral disagreements rest on different opinions about the facts or confusions of one or another sort...Differences in attitude and practice about these issues occur among people within the same larger society. There appears to me to be no objective way of settling these disagreements. That yields an argument for moral relativism that is similar for arguments for relativism about rest and motion, football, law, and language. (Harman 2012 p. 5)

Harman restricts his relativism to a particular class of moral judgments he calls "inner judgments." An inner judgment, for example, is 'she ought not to have stolen that necklace'. Inner judgments are different from other types of moral judgments in that they suppose a shared set of reasons and motivations between the person who is being judged and the person making the judgment.

We make inner judgments about a person only if we suppose that he is capable of being motivated by the relevant moral considerations. We make other sorts of judgment about those who we suppose are not susceptible of such motivation (Harman 1975 p. 4).

The truth of inner moral judgments is relative to social conventions. These types of moral facts are facts about subjective reasons; reasons that have their source in the agent's aims and goals. In his words, "conflicting judgments can be equally correct or equally justified" (Harman 1982 p. 308). So for someone who holds a radically different moral view than ours, we might say that they are a monster, evil, or that what they did shouldn't have happened. But we cannot say that such a person ought not to have done what they did (Harman 1975, p. 5). This is because we cannot say that they have a reason not to do what they did. In this way, the truth of moral judgments, or their justification, is not absolute or universal, but is relative to the traditions, convictions, or practices of a group of persons (Harman 1977 p. 133). Moral facts are mind dependent, meaning, for Harman they are, facts that we create, in contrast to scientific facts, which are mind-independent. Moral authority and normative force are *relative* to some social group. Standards that are authoritative for a group or society are the one's members of that group have agreed to through negotiations or bargaining. Moral judgments are only valid for groups of people who have made such agreements (Harman 1975).

The objection to the objectivity of morality based on disagreement has been widely discussed; as a result realists and other objectivists have a number of persuasive ways of responding to this argument.

One of the most well-trod responses to arguments from disagreement points out that disagreement in belief cannot lead by itself lead us to conclude that there is no objective fact of the matter in a particular domain (Rachels 2007). The fact that people have different moral beliefs or attitudes does not itself tell us whether or not there is an objective fact of the matter in ethics, just as it does not in any other subject matter.

However, a sophisticated relativist can easily accommodate this objection by saying that this is an inference to the best explanation, and so it is not making a deductive argument from disagreement to a lack of objectivity.

The second common objectivist response is that arguments from disagreement to relativism about values confuse epistemological questions with metaphysical ones. Widespread disagreements on ethical issues may reveal an epistemological problem about access to moral facts, but whether or not there are objective moral facts is a metaphysical issue. It could be that objective moral facts exist, but we simply do not have a very reliable or accurate way of accessing them (Shafer-Landau 2003 p. 227).

A third response is that there is less disagreement than it might at first seem. In fact, relativists have the burden on their theories of being able to explain all the agreement on moral issues that do exist (Sinnott-Armstrong 2007 p. 162). Realists have a ready-made explanation of moral agreement; people who agree may be recognizing the same moral facts.

Fourth, realists can use “companions in guilt” (or “partners in crime”) arguments, pointing out that in many other disciplines and areas of knowledge there is persistent disagreement. Yet these do not lead us to question the objectivity of those areas of inquiry, or the presence of facts in these areas. For example, in statistics, Bayesians disagree with non-Bayesians and in theoretical physics, adherents of string theory disagree with theorists of loop quantum gravity theory (Archard 2011 p. 121). In public discourse about science, there is widespread and persistent disagreement between proponents of intelligent design and evolution. In medicine, physicians often disagree with each other about diagnosis or treatment plans (McConnell 1984). Finally, in the

philosophy of mind disagreements remain between, for example, functionalism and eliminative materialism. However, no one considers using this disagreement as support for antirealism or relativism about statistics, physics, science, medicine, or the nature of the mind.

Mackie anticipates this kind of response.

Disagreement on questions in history or biology or cosmology does not show that there are no objective issues in these fields for investigators to disagree about. But such scientific disagreement results from speculative inferences or explanatory hypothesis based on inadequate evidence, and it is hardly plausible to interpret moral disagreement in the same way (Mackie 1977 p. 36).

However, Mackie's response here does not establish why moral disagreement cannot fit the same model as scientific disagreement, other than suggesting that it is "hardly plausible" (Mackie 1977 p. 36).

Despite examples of widespread *agreement* in ethics, no one would deny that disagreements remain. Since the skeptical argument from disagreement is an inference to the best explanation, realists and objectivists must offer alternative explanations of the ethical disagreement that remains. First there has been a limited amount of time and resources invested into thinking through ethical issues. Compared with sciences very few people have worked full time on ethical issues and few resources have been put into working out moral positions, developing theories, and looking at applied issues (Parfit 1984 p. 453-454; Brink 1989 p. 207).³¹

Second, widespread moral disagreement may be the result of a variety of kinds of distorting Influences. For example, failures of rationality may explain some remaining moral disagreement. Self-interest, including preference for one's own group and

³¹ For example, the National Institutes of Health's *Ethical, Legal and Social Implications* project is the largest bioethics initiative ever funded by government. Yet, its budget is only 5% of the total National Human Genome Research Institute's extramural budget.

unwillingness to think that we or our friends are doing something morally wrong may also be a significant barrier to moral agreement (Loeb 1998 p. 283). For example, Peter Singer and others have offered powerful arguments for why we should eliminate the suffering of animals in factory farms (Singer 1995). Yet few people have changed their practices. This may be attributable to the distorting effects of self-interest in enjoying the cheap consumption of these products. Cognitive distortions, such as status quo bias, an unjustified preference for the current state of affairs may also provide an explanation for why many find cognitive enhancement unethical (Bostrom and Ord 2006).

Third, in contrast to the sciences, few people who are not professionally involved in ethics consider it necessary to get the input of experts when it comes to forming ethical judgments (Brink 1989 p. 207). Thus there is an unwillingness to defer to experts in ethics. This may help some intractable public disagreement. For example, although much philosophical work has been done on the moral justification for a system of universal health care, few non-academics consult this literature or consider these arguments when forming their views on health care reform policies.

Fourth, many persistent disagreements in ethics stem from disagreement on the non-moral facts; here I am construing non-moral facts in the broadest sense. Non-moral facts that contribute to persistent ethical disagreement include, for example, religious beliefs, the truth or falsity of naturalism in ethics, the consequences of various actions and policies, and metaphysical and conceptual issues, such as the nature of personhood (Loeb 1998 p. 283). For example, on the issue of the morality of abortion, some persistent disagreement can be reduced to disagreement about the existence of a deity or the health and psychological effects that abortion has on the woman.

Finally, moral realists and objectivists can argue that there is some moral indeterminacy. Intractable disagreements may be the result of “moral ties” and may not be resolvable; these could be the result of incommensurable values or conceptual vagueness, for example (Brink 1989 p. 202; Shafer-Landau 1994).

In counter-response to some of these responses to arguments from disagreement, relativists and antirealists argue that they are not only talking about real world disagreement. Rather than making an inference to the best explanation, relativist or antirealist give an *a priori* argument (Shafer Landau and Cuneo, 2006 p. 219) .They claim that even idealized agents who met all of the above mentioned constraints (free from cognitive distortions, free from bias, agreed on all of the nonmoral facts, etc.) still would not converge on answers to moral questions, even if they were perfectly knowledgeable and rational. Realists have responded by pointing out that unless there are independent grounds for thinking that agents who are idealized and under specific constraints would still not agree on moral matters, this is begging the question against objectivism.

In part, whether this objection is successful or not depends on what is built into the concept of rationality that one is using. That is, if one grants that the agents under the idealized conditions are fully rational, one might be using a substantive or procedural version of rationality (Shafer-Landau 2003 p. 226-227). A substantive conception of rationality would mean that the agents in question “will be possessed of all such truths” that is “truths about what reasons there are” (Shafer-Landau 2003 p. 226). A merely procedural conception of rationality on the other hand, involves something like being able to “flawlessly identify best means to adopted ends, to infallibly

infer entailed conclusions from valid arguments” (Shafer-Landau 2003 p. 226). Shafer-Landau points out that the argument from disagreement, in this case depends on being able to rule out substantive rationality.

Another way of making the point about faulty reasoning is for the realist to concede that it is possible, even given perfect reasoning that disagreement on some moral issues may remain. But this is because “[t] there can be a gap between epistemic accessibility and truth” (Shafer-Landau and Cuneo 2006 p. 218)

A related counter-response by relativists against objectivism is that there is no established method for settling a moral dispute and there are such methods in science (Bambrough 1969; Rachels 1998). However, objectivists also have ways to respond to this worry. Even if we have no method of settling these disputes, this assumes that we have to have a way of settling a dispute in order for there to be an objective matter of fact. Second, realists can question whether in other areas where there is persistent disagreement there is truly agreement on a method, for example, in the debate between creationists and evolutionists.

8. Conclusions

This chapter presented several versions of the queerness argument against moral realism and in favor of various kinds of moral skepticism and antirealism. The two strains of the queerness objection that I will focus on in the following chapters have to do with issues of the normative force of morality. The objection from normative force can be broken down into objections having to do with a motivational sense of

normativity (the internalism objection) and objections having to do with the authority of morality.

In the following chapter, I will present the ways that non-cognitivists, and other antirealists, rely on claims about the nature of moral motivation to underpin arguments against cognitivism and moral realism. That is, I will focus on the internalism objection. I will briefly outline several versions of internalism and the major objections to internalism. I will defend moral realism by undermining the evidence for the kind of close relationship between motivation and moral judgment that internalists contend exists. I will also present the alternative position, externalism, and the general view I defend, a pluralistic motivational judgment externalism, again, the view that moral judgment is not inherently or necessarily tied to moral motivation.

In Chapter 3, I will focus on a key realist response to the motivation objection. The counterexample is of the amoralist who simply does not care about morality, though he is able to make moral judgments. Both realists and antirealists need to be able to give a plausible interpretation of these agents, who are both imaginary, and seem to be real.

Chapter four focuses on the strain of the queerness objection that has to do with the authority of morality. I start by disentangling the internalism objection from the authority objection and attempt to understand what this objection could amount to, as it has been understood in several different ways. I focus on two formulations of the objection, one made by Korsgaard, the other by Derek Parfit. There are independent reasons to reject both of their positions and I provide those. But I also argue that the authority objection does not pose a special problem for naturalist moral realism. In fact,

it is just as much a problem, if not a greater problem for constructivism, non-naturalist moral realism, or quasi-realism. Views other than moral naturalist realism, such as constructivism and quasi-realism, claim to be able to capture the objectivity of ethics but they have difficulty doing so. Although a view like Parfit's can capture the objectivity of morality, it leaves morality, specifically reasons, mysterious and *sui generis*. I suggest a positive view of what naturalist moral realism can offer in terms of the authority of morality, morality's objectivity.

Chapter 2: The Internalism Objection

Internalist theories of moral motivation claim that there is a necessary, essential, or inherent connection between making a moral judgment and being motivated to act on that moral judgment. This characterization of the relationship between moral judgment and motivation is often taken to be a problem for moral realism. This is because internalism is a premise in one prominent argument against moral realism, what I have called the internalism objection to moral realism. (See Chapter 1, section 3)

According to the internalism objection, moral judgments (taken either as a certain kind of mental state, or a verbal expression of that mental state) necessarily, essentially, or inherently motivate an agent to act on them. Cognitivism, often thought to be a crucial component of realism, claims that moral judgments are beliefs. The argument continues that because beliefs cannot motivate by themselves (The Humean theory of motivation, See Chapter 1), and moral judgments necessarily, essentially, or inherently motivate, moral judgments cannot be beliefs. In sum, the internalism objection asserts that the moral psychology that moral realists either do, or must embrace, is incompatible with a widely accepted and plausible claim about the relationship between moral judgment and moral motivation. Therefore, moral realism must be false.

Noncognitivists and expressivists use the internalism argument as support for their rejection of cognitivism and realism (Ayer 1936/2012; Stevenson 1937; Hare 1952; Gibbard 1992; Blackburn 1993). For these internalists, moral judgments are conative states; they are mental states which are much closer to desires than to beliefs. Thus, noncognitivists and expressivists are able to accept the idea that beliefs are motivationally inert (the Humean theory of motivation) while claiming that moral

judgments necessarily, essentially, or inherently, motivate. Since conative mental states like desires have the power to motivate by themselves, noncognitivists and expressivists present their theories as offering solutions to the internalism objection.

However, not all internalists are noncognitivists or expressivists. Many defenders of internalism hold the view for independent reasons. Internalist positions have been articulated that are said to be compatible with realism, including the positions of Thomas Nagel, David Wiggins, John McDowell and David McNaughton. These internalists resolve the conflict between realism, cognitivism and internalism by proposing alternative theories of motivation, either by rejecting the the Humean theory of motivation, or espousing internalism about reasons rather than about motivation (Shafer-Landau 2000 p. 279-280). Michael Smith's internalism is special in that it is part of his wider project of reconciling the Humean theory of motivation and desires, cognitivism, and internalism (Smith 1995),

A surprising feature of this debate is how little attention is paid in the literature to explicitly defining what motivation is. The vague working definition that most writing on internalism and externalism seem to be relying on is roughly, motivation is the mental state that causes the agent to act; the mental state without which the agent would not act.

In order to defend cognitivism and moral realism against the internalism objection, I will employ two main strategies: undermining the considerations that are generally taken to weigh in favor of internalism and presenting a version of externalism, pluralistic externalism, which can meet all of the objections that are usually made against externalism by its opponents.

In order to employ these strategies, in this chapter I will first, canvass several prominent internalist positions and discuss the four central arguments in favor of internalism, the argument from phenomenology, the argument from oddness, the argument from connection, and the empirical arguments. I will then show that none of these arguments make a strong case for internalism. In responding to each of these arguments or considerations in favor of internalism, I will be building the case for pluralistic externalism, as each one of these responses shows a way in which externalism can provide just as good, if not a better explanation of the phenomenon in question than internalism can. Having established that the major considerations in favor of internalism fail, I will discuss an independent argument against internalism, the moral indifference argument in detail in Chapter 3.

In the final section of this chapter, I will set out three conditions which the best version of externalism should meet, the flexibility condition, the empirical evidence condition, and the neutrality condition. I will then present the two main versions of externalism that have been articulated in the literature and explain why I take a combination of two forms of externalism to be most persuasive. I will explain how pluralistic externalism can meet the three conditions better than either one of the independent versions of externalism alone.

1. Varieties of Internalism

As a naturalist moral realist, David Brink rejects internalism. However, he has helpfully categorized the varying species of internalism. The most important distinction to make between different kinds of internalism is between internalism about reasons

and internalism about motivation (Brink 1989 p. 26). Internalism about reasons is the view that making a moral judgment “necessarily provides the agent with a reason to perform the moral action” while internalism about motivation is the view that making a moral judgment necessarily or inherently provides the agent with a motivation for action (Brink 1989 p. 26; Shafer-Landau 2000 p. 268). Because reasons and motivations can come apart, there is no need to be both an externalist about motivation and about reasons.³² These positions are independent and internalism about reasons will be only minimally explored here.

Most versions of internalism share a commitment to some version of the following three claims. One, morality is necessarily motivational. Brink specifically adds that according to internalism, “it is the concept of morality which shows that moral considerations necessarily motivate” (Brink 1989 p. 28). Two, not only do moral judgments necessarily motivate, they *essentially* motivate (Zangwill 2008 p. 95). The subtle difference between necessarily motivating and essentially motivating is clarified by Zangwill:

Imagine that it were somehow necessary that everyone has moral desires. In that case, all moral beliefs would necessarily be motivating, without that being the *essence* of moral beliefs. However, if motivation is essential to moral beliefs, that would *explain* why moral beliefs are necessarily motivating (Zangwill 2008, p. 95).

The third claim that many forms of internalism share is a commitment to *a priori* justification for their claim about the necessary and essential connection between morality and motivation:

³² Sometimes this distinction is made in terms of “motivating or explaining reasons” and “normative or justifying reasons” (Darwall 1997 p. 307). Unless otherwise noted, I am using “reasons” to refer to normative or justificatory kind, and motives to refer to what some would call motivational or explanatory reasons.

...this claim about the motivational power...of morality must be *a priori*. Since it is the concept of morality itself which determines this fact, the...motivational power of moral considerations cannot depend upon what the content of morality turns out to be, [or] facts about agents... (Brink 1986 p. 28).

There are a few who call themselves internalists but would not agree with this *a priori* characterization. For example, Jesse Prinz's version of internalism is based on his interpretation of empirical evidence and would thus likely reject all three of these claims (Prinz 2007). Later in this chapter I discuss some objections to thinking of the debate between internalism and externalism in terms of the *a priori* or conceptual analysis.

Claiming that morality necessarily motivates is still far too vague a characterization of internalism. Brink distinguishes three kinds of internalism (Brink 1989 p. 37). Agent internalism is the view that it is a conceptual truth that moral properties, facts, or obligations give the agent a motivation to act whether or not the agent recognizes the moral properties, facts or obligations (Brink 1989 p. 40).³³ Although agent internalism is seldom embraced, it is important to distinguish it from other forms of internalism because anti-realists, like Mackie, vacillate between the conceptions of internalism that they rely on to make the internalist objection against realism (Sinnott-Armstrong 2010 p. 59; Mackie 1977). Appraiser internalism is the position that it is a conceptual truth that a *moral belief* or *moral judgment* provides the appraiser with motivation for action" consistent with that belief or judgment (Brink 1989 p. 40). Unlike agent internalism, appraiser internalism claims that an agent's moral judgments, whether correct or not, motivate the agent (Brink 1989 p. 40). Thus the mental act of making a moral judgment is conceptually tied to having some motivation to act (Brink 1989 p. 41). Appraiser internalism is most common and most plausible, and

³³ Shafer-Landau calls this position "existence" internalism about motivation: "theories [that] claim that moral obligations necessarily motivate..." (Shafer-Landau 2000 p. 267).

unless specified otherwise, this is the kind of internalism to be discussed in this chapter. Finally, hybrid internalism says that it is a conceptual truth that when an agent makes a “correct or justified” moral judgment she is motivated (Brink 1989 p. 41). Mistaken moral judgments thus lack motivational power according to hybrid internalism.

There is a broad spectrum of strong and weak forms of internalism. Strong internalism says that if an agent genuinely makes a moral judgment, she will act on it. A second, weaker kind of internalism says that making a moral judgment provides the agent with *prima facie* motivation to act on that judgment, which competes with the agent’s other motivations. Weaker versions of internalism are more defensible and unless otherwise noted, will be the focus of the internalism debate in this chapter.

Sigrún Svavarsdóttir makes another distinction between strong and weak motivational judgment internalism (Svavarsdóttir 2005). For Svavarsdóttir, the defining characteristic of internalism, strong or weak, is the claim that it necessary condition of having made a moral judgment that one is motivated by it. Strong internalism makes the claim that the moral judgment has motivational power on its own, without the aid of any other mental state,³⁴ while weak internalism claims that a “moral judgment is not correctly ascribed to an agent unless he has the relevant motivation” (Svavarsdóttir 2005 p. 164).

Another way internalism can be strong or weak is by making restrictions on the class of agents that internalism applies to. Christian Miller points out that on some versions of internalism the connection between making a moral judgment and having a motivation to act on it holds for all agents (Miller 2003). On weaker versions of internalism, however, this connection only applies to a certain class of agents, those

³⁴ See for example, Russ Shafer-Landau 2003.

who are not in altered psychological conditions like apathy, depression, or weakness of will, or those who are rational, good, virtuous, or normal, for example (Miller 2003) (Brink 1989; Stocker 1979; Björklund et al. 2012; Mele 1996)

This weaker version of internalism (sometimes called “conditional internalism”) is easier to defend (Björklund et al. 2012 p. 126). Fredrik Bjorklund et al. have identified three different kinds of conditionality. The first condition is that the agent must be “psychologically normal,” this rules out “deviations from the normal functioning of deliberation and action guidance” (Bjorklund et al. 2012 p. 127). The internalist views of Blackburn (1984), Gibbard (2003) and Timmons (1999 p. 140) include this condition.

The second condition is “practical rational[ity],” that is internalism does not apply to agents who are in “conditions of decreased rational control of their actions” (Björklund et al. 2012 p. 127). This includes the internalist views of Korsgaard (1992) and Smith (1995), which focus on practical rationality. The third condition is “moral perceptiv[ness];” these internalisms only apply to agents that “really grasp the moral properties of actions” (Bjorklund et al. 2012, p. 127). Other internalists who make use of this condition are John McDowell (1978), David McNaughton (1988), Davis Wiggins (1987), Thomas Nagel (1970), and Jonathan Dancy (1993).

The ways that a particular internalist theory restricts the range of agents that it applies too is often connected to how the theory handles a forceful objection, the moral indifference, or amorality objection to internalism (See Chapter 3). The objection suggests that it is either conceivable or actual that some people can make genuine moral judgments but fail to be motivated by those judgments. Many internalists, as we will see, respond to this objection by claiming that if such agents did exist, they would

be eliminated by one of the conditions set on internalism. In other words, the internalist' diagnosis of the morally indifferent, will often put them into an exclusionary category. For example, the internalist might say that a morally indifferent agent is must not be rational, and thus internalism does not apply. The practice of restricting the scope of cases and kinds of agents that internalism applies to will be discussed in Chapter 3 on the moral indifference objection to internalism. The most important version of internalism in the overall argument of this project is a moderately weak version of appraiser internalism.

2. The Case for Internalism

Although the debate about internalism and externalism is one of moral psychology, much of the defense of internalism has been conducted on an *a priori* basis, in terms of conceptual analysis (Bjorklund, et al. 2012, p. 126). Internalists aim at establishing the necessary conditions for making a moral judgment. The necessary condition that internalists focus on is the condition that the agent must be motivated to some extent to act on the moral judgment. The central claim of internalism is that there is a necessary connection between making a moral judgment and being motivated by it, that this connection is built into our concept of moral judgment, and built into the way we use of moral language. Thus this connection between moral judgment and motivation should serve as a constraint on all of our metaethical theorizing (Smith 1995 p. 188).

A small number of both externalists and internalists are now also increasingly discussing empirical or *a posteriori* support for and against the positions (Roskies 2003 p. 331-333; 2006, 2008; Prinz 2006 p. 38-39; Prinz 2007 p. 42; Cholbi 2011 p. 29-34;

2006; Björklund et al. 2012; Björnsson 2012, 1998; Bjömsson 2002). The empirical evidence that has been discussed comes largely from psychological literature on mental disorders and disabilities, such as anti-social personality disorder, cases of injury to a patient's the ventromedial prefrontal cortex, and autism. How one interprets this body of research may depend largely on the position that one already holds on internalism or externalism and sometimes opposing sides of the debate cite the same studies as evidence for their position. Empirical evidence regarding moral indifference and purported cases of real-life amoralists will not be discussed here, but will be discussed briefly in Chapter 3.

Robert Lockie helpfully organizes the central arguments for internalism into three categories: arguments from phenomenology, arguments from oddness, and arguments from connection (Lockie 1998 p. 16). The arguments in those three categories tend to be part of the mainstream, *a priori*, conceptual analysis wing of internalism. Here I will outline examples of each of these three arguments for internalism and offer responses to each.

2.1 Phenomenology Argument for Internalism and Response

The argument from phenomenology says “that by reflecting on our experience of moral belief we will be led to realize that such is essentially motivating” (Lockie 1998 p. 16). Lockie cites the following argument from Mark Platts to illustrate this argument:

[W]hen we consider perception of moral value, I think it clear that, if introspective phenomenology can be our guide, moral perceptions manifest a unity of such a kind that potential motivation is indeed internal to them (Platts 1980 p. 81).

One way to respond to the argument from phenomenology is just to point out that one's own phenomenology does not support internalism. The fact that philosophers seem to have intractably different intuitions about this very issue has led some to argue that we actually possess different concepts of what a moral judgment is (Francén 2010).³⁵ But, as Copp points out, if internalists and externalists are using different concepts of moral obligation or moral judgment, then they are not actually disagreeing with each other at all, instead "they would be making claims about different kinds of obligation" or different kinds of moral judgments (Copp 1995 p. 191fn).³⁶ It may be the case that internalists and externalists are talking past one another; but surely there can be genuine disagreement about what the nature of moral judgments actually is, and how moral judgments relate to motivation in our mental and moral lives. But if externalists and internalists have clashing phenomenologies, we have no basis on which to prioritize one set of introspective observations over the other, no criterion by which to judge whose phenomenological reports should count as evidence and whose should be dismissed as mistaken. If this is true, then we will need to appeal to other sorts of evidence about the nature of moral judgments and their relationship to motivation.

Since internalism claims that there is a necessary, inherent, or essential connection between moral judgment and motivation, the kind of phenomenological evidence that internalists would need to lend support to this view would be very hard to

³⁵ At that point one option is to embrace pluralism about moral concepts: "the view that different internalist and externalist theses correctly accounts for different people's concepts of moral opinions, respectively" (Francén 2010 p. 117).

³⁶ Copp suggests that this is what Falk (1947) seems to take from the intractable clash of intuitions about "moral obligation" between internalists and externalists, that is that they have "different concepts of moral obligation" (Copp 1995a p. 191).

come by. It is not clear what the internalist actually finds when they reflect on their own mental life. How it is possible to discern a necessary connection between judgment and motivation through introspection? Even if we all shared the same phenomenology, reporting the same experience of motivation always following making a moral judgment, this would still not lend support to the claim that there is a *necessary* connection between making a moral judgment and being motivated. Hume's discussion of causality is applicable here; it may be that internalists merely observe a "constant conjunction" between moral judgments and motivation, but not a necessary one (Hume 1739-40/2000). We could just as easily conclude that this connection is contingent, as externalists contend it is.

Finally, the phenomenology argument is not persuasive because of the empirical evidence suggesting the difficulties of using introspection as a way of finding out about our own mental states, especially our own motivations (Nisbett and Bellows 1977; Nisbett and Wilson 1977; Wilson 2009; Pronin and Kugler 2007; Arcuri et al. 2008; Wilson and Bar-Anan 2008).

Examples from psychology range from perceptual illusions, mistaken reports about when we are experiencing a painful stimulus, to research on bias, choice-blindness, and confabulation. The studies that cast doubt on our ability to have accurate introspective insight into our own mental states are diverse and wide ranging and cannot be discussed in any detail here; a few brief examples suffice to make the point.

Nisbett and Wilson's widely cited study asked participants to select from five pairs of stockings which pair they thought was the best quality, although the stockings were virtually all the same, participants consistently chose the stockings that were on

the right side of the display. But when interviewed about why they chose the pair they did, they were not aware of the influence of where on the rack the stockings were placed and gave other reasons for their decision (Nisbett and Wilson 1977). From these and other similar results, Nisbett and Wilson concluded that we are often not aware of what causes us to behave or decide in one manner or another.

Choice blindness effects suggest that people confabulate to make sense of their decisions, even when in cases when they are misled about what decisions they actually made (Johansson et al. 2005; Johansson et al. 2006). In one experiment, participants were shown photos of women and asked to decide which woman was most attractive. In half of the cases, they were later shown the photo that they did not select and were misleadingly asked to explain why they chose that photo. The deception was only noticed in 28% of the participants (Johansson et al. 2006 p. 690). Even more interesting, the deceived participants explained with the same specificity and conviction as the non-deceived participants why they chose one woman instead of the other. There is also evidence that even though people may report being undecided about a particular political issue or decision, researchers can accurately predict the decisions they will make based on a test of their “implicit attitudes” (Arcuri et al. 2008; Galdi, Arcuri, and Gawronski 2008). Mentioning this research is meant to lessen the confidence that we should have in internalists’ phenomenological arguments for the necessary connection between moral judgments and motivation. If it is as difficult as this evidence suggests that it is for us to have transparent access to our own motivations and reasons when it comes to selecting a more attractive image or pair of hosiery, we likely have just as difficult a time, if not a more difficult time understanding our

motivations when it comes to such complex phenomenon as moral judgment and motivation.

2.2 Oddness Argument for Internalism and Responses

It seems strange if someone thinks or says that something is the right thing to do, but then lacks all motivation to do it, or wonders why they should do it at all (Lockie 1998 p. 16). This observation is the oddness argument for internalism. This argument can involve either the claim that the above mentioned scenario is *linguistically* odd or that it is *psychologically* odd. Mackie, for one, in the context of discussing Hume's argument that morality is not based on reason writes:

... it is *linguistically* [italics added] odd to use words like 'right' and 'wrong' with no prescriptive force-to say, for example 'x is right and Y is wrong, but of course it is entirely up to you whether you prefer what is right to what is wrong' (Mackie 1977 p. 55).

Linguistic oddness is often a precursor to the claim that such a person who speaks in this way must not understand the concept of right and wrong, or the concept of morality, because if they were using these terms correctly they would have to be motivated to some extent by their moral judgments. In other words, it is part of our practice of using moral terms that they carry "prescriptive force," in Mackie's language (Ibid). Similarly, Mackie asserts that "objective moral values" have "a power when known, automatically to influence the will" (Mackie 1977 p. 40).

An objective good would be sought by anyone who was acquainted with it, not because of any contingent fact that this person, or every person, is so constituted that they desire this end, but just because the end has to-be-pursuedness somehow built into it. Similarly, if there were objective principles of right and wrong, and wrong (possible) course of actions would have not-to-be-doneness somehow built into it (Mackie 1977 p. 96-97).

Mackie uses this as part of his larger argument for error theory and the denial of moral facts.

In his discussion of the meaning of the term 'good,' Stevenson sets out several "requirements, which appeal strongly to our common sense" that a definition of 'good' must meet (Stevenson 1937 p. 16). One of these requirements is that the "magnetism" of good must be captured in its definition: "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 had" (Stevenson 1937 p. 16). Blackburn makes a related, but weaker claim, when he says that he takes it as a "conceptual truth [that] to regard something as good is to feel a pull towards promoting or choosing it" (Blackburn 1984 p. 188).³⁷ Horgan and Timmons argue using a twin earth thought experiment, that we would not count a word as translatable as "good" unless the users of the word are typically motivated by the judgments they make about what is good and bad (Horgan and Timmons 1992, 1992).

The second type of oddness argument has to do with psychological oddness. The idea of psychological oddness is that it strikes us as an odd psychological tendency to be able to judge something morally right and then not be motivated to do it. For example, Dancy explains that what supports his internalism is the "'practicality' of morality...it would be odd for someone to say 'This action is wrong but I don't see that as at all relevant to my choice'" (Dancy 1993 p. 4). The phrase "relevant to my choice" can be read in two ways here. It may be a claim about motivation, that the agent in question is not motivated by his judgment that 'this action is wrong.' Or it may be a claim

³⁷ As a result, it is *prima facie* implausible, according to Blackburn, that moral judgments are expressions of beliefs.

about the authority of morality, in other words the authoritative significance or normative force of moral considerations.

One explanation for why Dancy and others take this agent's utterance to be so odd is this ambiguity between having no motivation and not recognizing the authoritative significance of moral considerations. Some of the appeal of the oddness argument may come from confusing the issue of moral motivation with the issue of authority. I have argued in Chapter 1 that motivational power of morality and moral judgment is distinct from the authority of morality and moral judgment. It is indeed odd to imagine someone who does not experience the authority, the categorical force of their own moral judgments in the way that most of us do. If we take the authoritativeness or morality to be distinct from the way that our moral judgments motivate us, then the oddness of a person or that person's words when they explain that they recognize that something is the right thing to do, but do not "see it as relevant to their choice" may only give us evidence that some kind of authority is built into our concept of moral judgment not necessarily that motivational power is built in. This issue will come up again when discussing the moral indifference argument against internalism (See Chapter 3).

A related, response is found in Svavarsdóttir, who suggests several reasons why internalists may have the intuitions that they have about such oddness (Svavarsdóttir 1999 p. 183). First, she suggests that internalists may simply have "a deep moral commitment that makes it hard for the individual in question to imagine how anyone could be motivationally unaffected by his moral judgments" (Svavarsdóttir 1999 p. 183). Their intuitions may also come from the "...optimism of the overzealous moralist that moral motivation is somehow guaranteed if we get people to see moral matters aright"

(Svavarsdóttir 1999 p. 183). This is the same kind of optimism found in Plato's claim that to know the good is to do the good. The oddness or "puzzlement," that the internalist detects, may be akin to the puzzlement one may experience when one encounters someone who is "...not moved by the beauty of a landscape;" it may feel as if they are not moved by it they could not possibly be seeing the same thing that you are (Svavarsdóttir 1999 p. 185).

Smith also relies on the oddness argument, claiming that we would be surprised if someone (for Smith, a good and strong-willed agent) espouses a moral view that forbids an action, and then performs that action (Smith 2004, p. 85). If a person continuously makes moral judgments and then acts contrary to them, we might wonder whether they are sincerely making the moral judgments at all (Smith 2004).

The observation of linguistic or psychological oddness is not one that the externalist needs to deny. Although, an externalist might deny that any of these situations strike them as odd. This is similar to the denial of the phenomenology of moral judgment and motivation. This move, however, leaves us only with a conflict of senses of oddness and won't get the debate very far; rather it may lead to merely endlessly shifting the burden of proof back and forth between the two sides. After all, the charge of oddness (or queerness, for that matter) can only be the very beginning of an argument; it cannot be the whole argument.

The question then becomes what we can make of this oddness. One strategy for responding to the oddness argument is to point out, as Lockie does, that an unusual connection between moral judgment and motivation, one in which an agent says that she judges an action morally wrong and then repeatedly performs it, or says she has no

accompanying motivation not to perform it (and in fact has no such motivation) is not the same as a “logically odd” or “incoherent” (Lockie 1998, p. 18). Perhaps this kind of relationship between moral judgment and motivation is merely statistically rare, and thus odd in that sense.

The internalist may respond that while the oddness argument does not fully support the claim that the very concept of a moral judgment involves motivation, externalism will have a harder time explaining why the rare, morally indifferent agent seems odd to us than internalism will. Internalists can claim that the reason such agents with an unusual connection between moral judgments and motivation seem odd, is that it is built into our concept of a moral judgment that it motivates to some extent. But if the externalist can offer a simple and straightforward explanation of oddness, then we do not yet have a reason to restrict the concept of moral judgment in this way.

If the internalists are claiming that it is logically odd or conceptually incoherent to claim that someone may make a moral judgment and yet fail to be motivated by it, they bear a very heavy burden of proof. One cannot simply assert that something is conceptually incoherent, one must show how this is the case. Here I reiterate a general criticism of the debate made by Nick Zangwill, who argues that while internalists often characterize their position in terms of a conceptual truth about the nature of moral judgment, questions about whether and the extent to which moral judgments motivate are psychological and empirical (Zangwill 2008 p. 99). I agree with Zangwill when he writes that his “view is that the issue between the various views of moral motivation is primarily a causal-explanatory one; it turns on the causal origin of certain motivations and actions” (Zangwill 2008 p. 99). Zangwill points out that even if we discovered that

motivational judgment internalism is in fact a conceptual truth, it would still be an open question whether or not people ever actually make the kind of judgments that our conceptual analysis requires (Zangwill 2008 p. 100).

Further, the empirical evidence that does exist favors motivational externalism (Zangwill 2008 p. 99). I side with Zangwill in rejecting the method of pure conceptual analysis for understanding the relationship between judgment and motivation.

The externalist has several responses available to the oddness argument for internalism. If we should interpret the oddness argument to be about the odd psychological makeup that a person would have to have in order to be able to make a moral judgment and fail to be motivated by it, the externalist can simply deny that this strikes her as oddness altogether. A more promising externalist strategy is to offer an explanation of why this is an odd (perhaps in reality, rare) but conceptually possible way for a moral judgment to function in one's mental economy. Finally, the externalist can point out the ways in which oddness is deflated when we remove moral authority and moral optimism from the observation.

2.3 The Connection Argument and Responses

The strongest and most widely discussed argument for internalism is the argument from connection. The argument from connection says that "we must explain the fact, so obvious that it is easily taken for granted, that moral beliefs are commonly and non-coincidentally *connected* with moral action" (Lockie 1998 p. 17). This predictable and regular relationship between moral judgments, motivation, and actions is what Smith thinks only internalism can explain and he has given one of the most

sustained presentations and defenses of this argument for internalism (Lockie, 1998 p. 16). The idea is that the common sense observation that there is a regular and reliable connection between making a moral judgment and being motivated by that judgment is easily explained by internalism, but for various reasons is difficult or impossible for externalism to adequately explain. Smith's connection argument is one that aims to shift an explanatory burden onto externalism.

Rather than looking at connection arguments in general, it makes more sense to focus on Smith's connection argument in detail, since his is the most widely discussed and since he offers a restricted version of internalism it is most plausible. In order to explain Smith's argument for internalism and related objection to externalism, I will have to provide a summary of part of Smith's project in the *Moral Problem*. Smith makes the following supposedly intuitive observation about the nature of our moral lives:

[b]y all accounts, it is a striking fact about moral motivation that a change in motivation follows reliably in the wake of a change in moral judgment, at least in the good and strong-willed person (Smith 2004 p. 71).

Smith's discussion of internalism is restricted to a particular type of agent, one who is good (sometimes virtuous) and strong-willed (thus it is a weaker conditional form of appraiser internalism). When such a person changes their moral judgment from the judgment that 'x is right,' to the judgment that 'x is wrong,' his or her motivation and tendency to act in accordance with his or her judgment also changes in a predictable way (Smith 2004 p. 74). For example, if I begin by judging that giving to charity is not morally required, I will have no motivation to give to charity, barring motivation to do so from some other source, like the desire to get a tax break. If, after reading Peter Singer's "Famine Affluence and Morality" I come to make a different judgment, namely

that it is morally required that I give to charity, I will be motivated to give. He argues that this reliable connection between making a moral judgment and being motivated by it can be explained in one of three ways:

On the one hand we can say that the reliable connection between judgment and motivation is to be explained *internally*: it follows directly from the content of the moral judgment itself. The idea will then be either that the belief that an act is right produces a corresponding motivation (this is the rationalists alternative), or perhaps that the attitude of accepting that an act is right is itself identical with the state of being motivated (this is the expressivist'). Or on the other hand, we can say that the reliable connection between judgment and motivation is to be explained *externally*: it follows from the content of the motivational dispositions possessed by the good and strong-willed person (Smith 2004 p. 72) [italics added].

The three ways of explaining the reliable connection are, according to Smith; one, the content of the moral judgment produces a motivation (internalism, rationalism), two that the judgment is an affective state and thus already a motivational state (internalism, expressivism), or three, that the judgment and the motivation are separate (externalism). The solution that Smith ultimately defends is the internalist rationalist position, explanation one. The expressivist explanation, explanation two, is not an acceptable option for Smith because it is in direct conflict with the objectivism that he seeks to defend elsewhere. Importantly for this chapter, Smith goes on to argue that there are also serious defects with the externalist explanations, explanation three.

Smith argues that only theories that embrace what he calls the “practicality requirement” can explain the reliability of this connection in the “good and strong willed person” (Smith 2004 p. 72). The practicality requirement, put in formal terms by Copp, says: “It is necessarily the case that a person who believes she is morally required to do Φ in circumstances C is motivated to do Φ in C unless she is ‘practically irrational’” (Copp 1997 p. 36). More specifically, the “desire [to do Φ] is caused by the belief that it

would be right to do the thing” (Copp 1997 p. 36). The practicality requirement is a conceptual truth and a necessary truth about moral judgment.

Crucially for Smith’s internalism, the desire that is at work in his practicality requirement is the desire to do the right thing in a *direct* and *un-derived* sense. Smith calls this a *de re* desire to do the right thing, which is caused necessarily in a person who judges that something is the right thing to do. A *de re* desire to do the right is a desire for example, to help a friend, or to relieve suffering, or to tell truth. In contrast, a *de dicto* desire to do the right thing, is a desire simply to do what is right, the desire then to help a friend, or to relieve suffering, or to tell truth, is derived from this desire and is an indirect desire. This distinction between a *de re* desire to do the right thing and a *de dicto* desire to do the right thing is crucial for his argument.

Smith charges that the externalist is unable to adequately explain why our motivation changes as our moral judgments change. Smith considers what he takes to be the externalists’ only competing explanation of the reliable connection between moral judgment and motivation (Smith 2004 p. 74). All that the externalist can do, Smith argues is claim that the reliable connection between judgment and motivation in “the good person” is the result of a “*de dicto* desire to do-whatever-is-morally-required” (Copp 1997 p. 49). This desire is indirect and derived, unlike the desires that the good and strong-willed person has according to his internalism. So Smith thinks that the externalist must posit the existence of a belief-desire pair to explain a shift in motivation that occurs with a shift in moral judgment. This explanation fails, however, according to Smith, because it means that the externalist agent is engaging in a kind of “moral fetishism” (Smith 2004 p. 128). I will refer to this as the “fetishism argument”:

the good person is, at bottom, motivated to do what is right, where this is read *de dicto* and not *de re*, and that is really a quite implausible claim. 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, their fellows, people getting what they deserve, justice, equality and the like, not just one thing: doing what they believe to be right, where this is read *de dicto* and not *de re*. Indeed common sense tells us that being so motivated is a fetish or a moral vice, not the one and only moral virtue (Smith 200, p. 75) [italics added].

The good externalist agent who is motivated by a desire to be moral (the *de dicto* desire to do what is right) cannot possibly be morally praiseworthy. Smith thinks that even the best morally good externalist agent cannot actually care about the objects of her moral judgments; she does not care about being honest or compassionate. Nor does she care about the people that are affected by her actions. Instead she cares about “doing the right thing” (read *de dicto* and not *de re*, in Smith’s language) (Smith 2004 p. 83). As Shafer-Landau puts it, “we nowadays frown on the view that sees those who assiduously attend to the moral qualities of their conduct as paragons of virtue...those possessed of the motive of duty seem just a tad too preoccupied with the moral worth of their actions” (Shafer-Landau 2000 p. 285).

Smith concludes that the only avenue open to an externalist to explain the reliable connection between moral judgment and motivation is inconsistent with common sense about the nature of a good and strong-willed person. And since the expressivist explanation of the connection is inconsistent with Smith’s other theoretical commitments, he concludes that his unique form of internalism, captured by the practicality requirement, is the only acceptable alternative.

Externalists have three main strategies for responding to the connection argument. First, the externalist may challenge the existence of the reliable connection

between moral judgment and motivation. This response usually involves counter-examples, pointing out cases where the connection does not hold, as in cases of moral indifference and amorality (See Chapter 3). A small part of this response will be developed in section 2.3.1. The second option for an externalist is to show how an externalist need not adopt the *de dicto* desire view as their theory of moral motivation. Several externalists have taken on this project; it is part of the wider effort to articulate the most plausible version of externalism. These arguments show why we need not believe that according to externalism agents have only one kind of moral motivation (*de dicto* desires). This will be done in section 2.3.2. The third avenue for defending externalism is to show that even if the *de dicto* desire view is the only option available to externalists, it is not problematic or inconsistent with our view of the good agent, to address the moral fetishism argument. Instead, *de dicto* desires can play some role in the good agents' moral psychology. This is argued for in section 2.3.3.

There has been a vast amount of discussion in the metaethical literature on the argument from connection, responses to it, and Smith's account of moral motivation. Every response and counter-response cannot be accounted for here. In the following sections I will present the objections to the connection argument that are salient to the species of externalism that I find most persuasive, pluralistic externalism. This allows me to present a wide variety of responses to the argument from connection. This is because the kind of externalism I endorse is a fairly permissive, in that it includes the possibility of a wide variety of motivational mechanisms and sources for moral motivation.

2.3.1 There is No Reliable Connection

Smith stipulates that the reliable connection between making a moral judgment and being motivated only holds in agents who are both good and strong-willed. Obviously this makes Smith's version of internalism applicable to a fairly small group of moral agents. As Copp illustrates with the case of Carol, the reliable connection that Smith describes does not necessarily exist in an agent who is only good, but not strong-willed (Copp 1995 p. 209).

Suppose a demagogue convinces Carol that she is in fact obligated to support capital punishment. A while ago you persuaded her that she ought to oppose it, and you managed to link her fundamentally compassionate nature to her opposition to capital punishment. The link is cemented so firmly, let us suppose, that the demagogue is unable to overcome Carol's revulsion to the death penalty. He successfully convinces her that she is obligated to support it, but she continues to oppose it. Now Carol may well be a good person, it seems to me, despite the fact that she is weak-willed; her belief about capital punishment has changed without an accompanying change in motivation, yet (I assume) her fundamental motivations or values are appropriate. Indeed, her goodness seems to depend on her not having the desire to-do-whatever-she-is-obligated-to-do, for if she has this desire, her wavering beliefs about her obligations would bring about wavering motivations (Copp 1995a p. 211).

If an agent is good, but weak-willed, she may lack the reliable connection between moral judgment and motivation. In response, Smith might disagree with Copp's claim that the demagogue has truly "convinced" Carol that she is "obligated" to support the death penalty. But if he accepts that aspect of the case, it seems Smith would have to say that Carol is not a good person at all. Smith defines what a good person is in the following way: "it is constitutive of being a morally good person that you have direct concern for what you think is right" (Smith 1994 p. 76).

So Copp's case is doing more than illustrating how a person who is good, but not strong willed, can fail to have the reliable connection. Instead, he is challenging Smith's notion of a good person. In the case of Carol, he assumes we will all agree that Carol is

good, as the source of one's goodness is to be found in "one's motivations rather than in one's beliefs or the connection between one's beliefs and one's motivations" (Copp 1995a, p. 211).³⁸ In offering an account of the relationship between moral judgment and motivation, I wish to remain neutral on the first order normative issue of what makes an agent a good agent as much as possible, be it their motivation, actions, character, or otherwise. This is a debate I will only engage in to the extent that I wish to leave this an open question, where other theorists, like Smith, unnecessarily close the question, and make certain assumptions about what a good agent has to be like. So while I do not endorse Copp's description of what makes Carol good, his case is useful in revealing the assumptions Smith is making about what makes an agent good.

Copp offers a second case that questions whether the connection between moral judgment and motivation always obtains even in agents that are *both* "good and strong-willed" (Copp 1997 p. 50). To illustrate this doubt, Copp imagines a case of a good and strong-willed person who comes to believe that abstaining from sex is morally required. This belief happens to be false; abstaining from sex is not morally required. Despite the fact that this person:

values doing all the other things she believes she is required to do, and all her other (basic) beliefs about what she is actually required to do are accurate...she desires as any good person would desire, this agent never comes to see abstaining as being valuable, and never comes to desire to abstain (Copp 1997 p. 51fn.).

In this case, it seems that the connection between the moral judgment that she makes and her motivation has been severed, despite the fact that she is good, strong willed,

³⁸ Here Copp's unique "standard based account" of moral obligation and moral judgment is important. Whether the agent is good or not is a matter of whether or not she "subscribes to the appropriate moral standards", which Carol does, despite the fact that she has the mistaken moral belief that she is obligated to support the capital punishment. Most of the time one's moral beliefs and the standards one subscribes to are lined up, but they do not necessarily have to be (Copp 1995a p. 211).

and lacks practical irrationality.³⁹ Smith might respond to this example by saying that we cannot attribute to her the judgment that “abstaining from sex is morally required” in light of these other facts about her desires and what she values. But if he makes such a claim, he seems to be assuming part of what his argument is trying to establish, since the case description of Carol doesn’t sound incoherent or implausible.

In the following chapter I will discuss a related, but wider range of phenomenon that I am calling forms of “moral indifference” (Zangwill 2008). Cases of moral indifference are a further way of challenging the reliable connection between moral judgment and motivation. I will suggest (with Zangwill and others) that moral indifference is not only present in pathological personality types, like psychopaths, but is also a common phenomenon among everyday moral agents like you and me. And indifference thus, serves as a strong counter-example to not only Smith’s argument, but the general form of the reliable connection argument.

2.3.2 Alternatives to *De Dicto* Desires

An important part of Smith’s argument against externalism is that externalists’ only way of explaining the reliable connection between moral judgment and motivation is to posit a *de dicto* desire to the morally right thing (Smith 1995). Smith argues that this puts externalists in the position of having to embrace moral fetishism.

Externalists respond with several persuasive arguments that the *de dicto* desire account that Smith presents is not the only possible externalist account of the motivating state that causes the good person (or any person) to act morally.

³⁹ Cuneo also briefly questions the reliability of the connection, suggesting that “one could have some slight change in moral judgment and experience no change in moral motivation” (Cuneo 1999 p. 371fn.).

The first suggestion that several externalists have made is that it is plausible to think that *along with de dicto* desires to do what is right, good and strong-willed agents possess a host of other desires that play a motivational role as well. The pluralistic externalism that I find most persuasive embraces the following defense of a multitude of desires and affective states that might be the motivation that gets a person to act on their moral judgments.

Several externalists have pointed out that we have no reason to think that people have one single motivation for any particular morally good act (here an act includes refraining from an action) (Cuneo 1999; Svavarsdóttir 2005; Shafer-Landau 2000; Copp 1997; Olson 2002).

A morally good and strong willed person may have a *de dicto* desire to do the right thing (which Svavarsdóttir calls the “desire to be moral” and Shafer-Landau calls “the motive of duty”) and at the same time possess many other morally relevant “direct desires,” for example, the desire for her friends and family to be happy, the desire for world peace, the desire to be kind, the desire to be a good friend, or the desire to not hurt people, etc. (Copp 1997 p. 50; Cuneo 1999 p. 370; Shafer-Landau 2003 p. 158, 285). Olson also points out the possibility of “motivational over-determination” (Olson 2002 p. 91).

All of these desires can remain external to our moral judgment; in that just because we make a moral judgment does not necessitate that we have one of these direct desires to act on it. In these cases there is no need to introduce the *de dicto* desire to the right thing to explain the reliability of the connection between moral judgment and motivation.

It sounds plausible to suggest, as externalists do, that the reason our moral judgments often motivate us to some extent to act on them is that our moral judgments connect up in the relevant way with related desires we have.⁴⁰ It is important to be clear about how this suggestion differs from internalist accounts of motivation. The abovementioned suggestions are not the claiming, as internalists do, that once we have formed a moral judgment, either the moral judgment by itself motivates us to act, or the moral judgment necessarily causes a desire that motivates us to act.

At the same time this externalist account is not ruling out the possibility that sometimes we *do* act on the *de dicto* desire to do the right thing. What is appealing about this kind of “motivational over-determination” is that it seems to capture the concern that someone like Smith has about good and strong willed moral agents being motivated directly by the objects of moral concern, justice, friendship, and so on. On this version of externalist explanation this is often the case. But it is not because the moral judgment necessarily motivates. It is because good and strong-willed agents possess a host of other desires to promote justice, friendship, etc. It is unlikely that agents that are not good or strong-willed also possess many of these desires.

Another advantage of this externalist position is that if someone lacks the desire with the appropriate content (desire to promote the welfare of the poor, for example) and lacks the *de dicto* desire to do the right thing, then it is possible that even if he makes the moral judgment that he ought to help the poor, he may lack any motivation to help the poor. Thus this kind of pluralism about motivation is consistent with the phenomenon of moral indifference.

⁴⁰ How these desires are created is not being addressed here.

According to Smith however, the presence of any *de dicto* desire to do the right thing, even if it is only *one* motivating desire among many, somehow taints the goodness of the moral agent or the value of the actions they perform (Olson 2002 p. 91).

To illustrate how odd this result of Smith's theory is, Copp considers this case: "[i]f he (Smith) is correct, then no (entirely) good person has this desire [the *de dicto* desire to do what is right], not even a good person who is also strong-willed" (Copp 1995 p. 212). The point is that Smith's account implies that even if a strong-willed person has a whole host of *de re*, "direct concerns" for morally relevant things, such as "justice, honesty, and the well-being of her friends," if even one *de dicto* desire is added to her moral psychology, she no longer qualifies as a good agent (Copp 1995a p. 212). This seems to be inconsistent with our beliefs about morally good agents. Simply acquiring a new *de dicto* desire, while all of an agent's previous direct concerns remain intact, cannot be enough to no longer call her a good agent (Copp 1995a p. 212).⁴¹

The views that I have just discussed all suggest that we have no reason to think that an externalist has to say that the good and strong-willed agent (or any agent) has only one kind of motivation or desire that causes them to do the right thing or be inclined to do so. Further, that if the *de dicto* desire to do the right thing, is one among many

⁴¹ Copp explains that Smith would probably want to know, in this case, "what would happen to Dena if she were to acquire a new belief about what she is obligated to do", to test the reliability of the connection between judgment and motivation (Copp 1995a p. 212). Would she be motivated by the *de dicto* desire "to-do-whatever-she-is-obligated-to-do" or would she have a direct, *de re* concern for the object of her new moral obligation? Copp suggests that this might be the way that this agent acquires a new "desire to do a new thing," her desire or interest in the object of the new moral obligation could be derivative. It is very strange to say that this now makes Dena a bad person, or no longer a good person (Copp 1995a p. 212). "Nor does it seem that Dena would have been a better person if she had been able immediately to acquire a direct and non-derivative desire to do the new thing" (Copp 1995a p. 212-13).

other morally relevant desires in the agent's motivational set, then it would be odd for this to cast any doubt on the agent's goodness.

The next group of arguments provides further reasons to think that *de dicto* desires are not the only explanation for the reliable connection between moral judgment and motivation that externalists have at their disposal. First, externalists need to use the *de dicto* desire to do the right thing to explain the reliable connection in only a very limited number of cases. Second, I make the case that there are other theories of moral motivation that do not involve the *de dicto* desire view at all and are explanatorily adequate.

Strandberg defends externalism by arguing that it does not necessarily have to resort to a *de dicto* desire view to explain the reliable connection in the vast majority of cases. Strandberg sees Smith's fetishism argument as posing the following question for the externalist: "What is the *content* of the motivational states that constitute moral motivation? [emphasis added]" (Strandberg 2007 p. 250). Strandberg restates the difference between the *de re* and *de dicto* desires in terms of the propositional content of the desires: "[i]f a person is motivated by a desire *de dicto* to do what she judges to be right, her desire has a content that involves the concept of rightness; the concept of rightness figures as part of the intentional content of her desire" (Strandberg 2007 p. 251). This is exactly what Smith finds objectionable. While the person who is "motivated by a 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 p. 251). That is the account that Smith and internalists in general are able to give.

Strandberg offers the following account of one way an externalist can avoid having to endorse *de dicto* desires view:

Assume first that the person in question believes that certain features make actions right. We may for the sake of simplicity call these features F. To illustrate we may suppose that F consists, among other things, in helping people in need. Next we may assume she has a desire to perform actions that have these features F. Externalists can now account for the reliable connection at issue by assuming that the person in question alters her beliefs as regards what actions have F (Strandberg 2007 p. 253).

The agent has a belief about what one of the right making features of actions is, that is, for example, that they help people. What the agent in question is motivated by is the *de re*, unmediated desire to help people. At one time she may not be motivated to give to the homeless because she believes it is not a way to help them, and thus the action lacks the crucial right making feature. She may believe that it is actually be a way of harming them, and thus has a wrong making feature instead. If, at a later time, she changes her mind and comes to believe that giving to the homeless actually does help them, she will then judge that it is right to give money to the homeless, and she is now motivated to do this. Now that she believes that giving money to them helps them, she is motivated to give. This account does not involve the agent ever being motivated by a *de dicto* desire to do the right thing. Instead, it is “a desire *de re* to do what is judged to be right;” she desires to help people and judges that it is the right thing to do (Strandberg 2007 p. 253). In this case Smith can have no complaint of fetishism.⁴²

Cuneo makes a related argument; the externalist can analyze the source of the reliable connection in terms of the concerns that a virtuous person has:

⁴² Similarly, Cuneo argues that externalists can explain the reliable connection in terms of “a change in belief” about whether or not a particular action in fact promotes one of the goods she has a *de re* desire for and is thus motivated by, in Cuneo’s example it is human flourishing (Cuneo 1999 p. 372).

The virtuous person possesses a whole constellation of virtues, and hence, a whole pack of correlative concerns. Thus, when some strong-willed virtuous agent judges that one or another action is morally obligatory, helpful, compassionate, etc., that virtuous person is typically motivated to some appropriate degree or other to take that action because she typically possesses the requisite concerns with which the content of that judgment can combine in the right fashion (Cuneo 1999 p. 370).

Virtuous agents have standing concerns about helpfulness, being a good friend, compassion, and the like, and when they judge that an action has those features, the features themselves motivate the agent to act accordingly (Cuneo 1999 p. 370).

The objection of an internalist like Smith to these accounts would be that if the agent changes her mind about what the right making features of actions are, then her desire to do actions with that feature may not track her change in belief (Strandberg 2007 p. 254).

Strandberg addresses this objection, arguing the agent may have already had a “pre-existing” desire “to perform actions which have these features,” that is the new “right-making” feature (Strandberg 2007 p. 255). The agent may believe that the right making feature of actions is that they respect other people’s autonomy. This same agent desires to help people and desires to respect people’s autonomy. She believes that giving money to the homeless is a way of helping people but is inconsistent with respecting their autonomy. So while she is somewhat motivated to give money to them, she has a stronger motivation to refrain. Over time, she comes to believe that what actually makes actions right (or what is more important) is that they help people, rather than that they respect their autonomy. She then judges that giving to the homeless is right and is also motivated all things considered to do so, because she had the pre-existing motivation to help people. Here externalists do not have to rely on any *de dicto*

desires with the content that includes the concept of rightness. If this is not the case however, and the agent had no pre-existing desire to help people, then externalists may only be able to explain her new desire to do what she now believes is right in terms of a *de dicto* desire (Strandberg 2007 p. 255).⁴³

Dreier proposes a second-order desires externalism, where the second order desire is the “desire to desire to do whatever is in fact right” (Dreier 2000 p. 623). Dreier explains the second order desire to desire, as being analogous to having a “maieutic end.” The idea is that people are capable of choosing their own final ends; “[a] final end is one that is pursued for its own sake and not merely for the sake of something else” (Dreier 2000 p. 629). This is possible because of a “maieutic end,” which is “an end achieved through the process of coming that have other ends” (Dreier 2000 p. 630). Dreier’s example of pursuing a career in medicine is helpful. If one desires to have a satisfying career, that is an instrumental end, but it is also a maieutic end. In order to achieve it you have to adopt some other ends as final ends in themselves:

[One] chooses medicine for this reason, an instrumental end. But once you have chosen medicine as a career, the goals you have-of relieving suffering, being a respected member of the medical community, and so forth-are goals you pursue for their own sakes. Having adopted these goals for one reason, you then pursue them not for that reason but for reasons internal to the outlook of your profession (Dreier 2000 p. 630).

Dreier likens maieutic ends to second order desires, “to have a maieutic end is to desire to have certain ends, or ends of a certain type” (Dreier 2000 p. 630). Applying this machinery to externalism and the moral case, Dreier argues that the agent who has a second order desire to “value for their own sake those things that are (or that one

⁴³ Of course, Smith will object that the agent can no longer be counted as good (Strandberg 2007, p. 255).

believes to be) morally right,” might fit with the criterion that Smith thinks a virtuous agent must meet (Dreier 2000 p. 630).

The two criterion or tests that Smith presents for any externalist model of motivation are the reliable connection and the avoidance of moral fetishism. Dreier argues that this kind of moral agent can meet both of those standards. First consider the reliable connection. If this agent’s (David’s) moral judgments change and he now believes he is morally obligated to vote for a different political party, he will experience a reliable change in motivation and likely a change in behavior as well:

David’s motivations will track his changes in moral view, for once he comes to believe that, say, voting for Social Democrats is right, David’s second order desire will kick in and he will actually desire to vote for Social Democrats (Dreier 2000 p. 630-31).

The charge of fetishism does not apply here. When Dreier’s imagined virtuous agent

...desires to vote for Social Democrats because (a genuinely causal ‘because’) he believes that it is right to do so, but once the desire is generated it is not conditioned on the belief that voting for Social Democrats is right, any more than out doctor’s desire to cure the sick is conditioned on his belief that medicine would be a satisfying career (Dreier 2000 p. 631).

Yet the generation of the desire is not necessarily connected to making the moral judgment that voting for the Social Democrats is right. So this view is distinct from internalism.

This section has provided some of the alternatives that externalists have to the pure *de dicto* desire view of moral motivation. These views are all fairly consistent with one another. If we take these accounts of externalist alternatives to *de dicto* desire motivation to be complimentary rather than competitors, the following picture emerges. While agents may have *de dicto* desires to do the right thing, these are not the only desires that motivate them to act on their moral judgments. Agents have a wide range

of direct *de re* desires that relate to their moral judgments. They may have desires for justice to be realized, the desire to be a good friend, the desire to end the suffering of others, or for their families to be happy. It is convenient to call these desires, but this could include a range of propositional attitude types, such as concerns for or caring about. So I may judge that it is the right thing to do to call my Mom on her birthday. I believe that calling her on her birthday will make her happy. And I am motivated to call my mom on her birthday by my desire for my mom to be happy.

This fits well with Strandberg's suggestion that agents make judgments about whether or not certain actions have "right making" features or properties. Agents sometimes judge that an act has a particular right making property, for example, that it will make someone happy. When we are motivated to act on this judgment it is often because we have a direct *de re* desire for those right making properties, we have a direct *de re* desire to make people happy.

Another way to think about the relationship between our desires and our moral judgments is Dreier's suggestion that we have a second order desire to desire whatever is morally right. These complimentary accounts of alternatives to the *de dicto* desire view add up to a pluralistic externalist picture of motivation that is far more subtle and has much more explanatory power than Smith and other internalists attribute to externalism.

2.3.3 Defending *De dicto* Desires and Embracing Moral Fetishism

It is important to defend *de dicto* desires because the ultimate account of externalism that I advocate is a pluralistic account. This view says that *de dicto* desires

can and do play a role in our moral psychology, whether or not we are good and strong-willed agents. The pluralistic account stipulates, however, that these are not the only desires involved.

I will remain ultimately neutral on what it is that makes a good agent morally good. But Smith engages in that debate when he claims that there is something morally wrong with “moral fetishism” and that the good agent cannot be motivated by *de dicto* desires. A fundamental problem with Smith’s charge of moral fetishism against *de dicto* desire views, is his assumption that “the moral desires which motivate a person to act on a moral belief would have to be what makes a person morally good” (Smith 1994 p. 74-76; Zangwill 2008 p. 148). Smith thinks that the person who is motivated to act on their moral judgments by the *de dicto* desire would fail to be a virtuous or morally good agent. But as Zangwill points out, why should we think that what motivates someone to do the right thing can make or break them as a good or virtuous agent? Smith would at least need a separate defense of that substantive moral claim. An agent’s “virtue” or goodness does not necessarily have to be grounded in the desires that motivate us to act on our moral judgments (Zangwill 2008 p. 148).

In an effort to keep open the possibilities of what kind of motivation (if any) determines whether or not an agent is good, I argue against Smith’s condemnation of moral fetishism. I argue instead that it is perfectly plausible and consistent with much common sense and philosophical thought about good agents to say that they may be motivated by *de dicto* desires to do the right thing.

A more appealing way of thinking of being motivated by a *de dicto* desire to do the right thing, is that such an agent is keeping an open mind to the content of morality

while remaining motivationally committed to acting on *whatever* morality requires. In fact this is a characteristic that Socrates' embodies in the early and middle Platonic dialogues, if we consider the Iakovos Vasiliou's contrast between "aiming principles" and "determining principles":

An aiming principle tells the agent what overall aim she ought to have in acting, for example, to do the virtuous action; because this particular aiming principle is so important to Socrates and Plato, I shall give it a name: 'the supremacy of virtue'...By contrast, a *determining principle* (e.g., a proposed moral rule such as "Never kill anyone") would be one that actually determines which action or action-type is forbidden or required; once you adopt "Never kill anyone" as a principle, then, at least as far as that principle is concerned, if an action involves killing someone, it is forbidden...According to Cleitophon, then while Socrates has a clear answer to an aiming question...he utterly fails to offer answers to determining questions" (Vasiliou 2008 p. 2-3)[italics added].

Socrates does not claim to have knowledge of determining principles; instead, he espouses knowledge and adherence to the aiming principle of the supremacy of virtue. In the *Crito*, for example, Socrates engages dialectically with Crito, determined to decide whether he should escape from prison or not by finding out which course of action is morally better (Plato 2002). Whatever course of action turns out to be the good action, is the one he will choose. On Vasiliou's analysis, Socrates, as a virtuous agent, is committed to the supremacy of virtue. This gives us insight into his motivation:

we know that the agent is motivated to determine what the virtuous or vicious action is, and base her acting (or not acting) on that ground. Thus when the agent committed to SV acts, she will refrain from acting viciously because the action is contrary to virtue and will act virtuously because the action is virtuous (Vasiliou 2008 p. 43).

In other words, when the virtuous agent does the virtuous thing, the agent is "not acting for a further end" (Vasiliou 2008 p. 44-5). Thus the virtuous agent's ultimate aim is to act virtuously. Eventually, Vasiliou gives an account of the source of this motivational

commitment to “put[ting] virtue above all”, one that I will not go into here (Vasiliou 2008, p. 44-5).

I cautiously suggest that one way to think of the aiming principle is akin to a *de dicto* desire to do the right thing. Smith charges that the person who is motivated to behave morally solely out of some kind of a desire to do the right thing is morally flawed. If my translation of aiming principle/supremacy of virtue into a *de dicto* desire to do the right thing is defensible, then Smith would have to reject Socrates as a virtuous agent, as well as Platonic notions of the relationship between virtue and motivation. Surely, the fact that Socrates, presumptively a paradigm of a moral agent, seems to be motivated in this way should lead us to question Smith’s assumption. My appeal to Socrates and Plato is not meant to be a mere appeal to authority to lend credence to this defense of externalism. But it does seem that there are perfectly plausible ways of understanding ideal moral agents that include motivation on the basis of *de dicto* desires to do what is right.

Olson also argues that there are situations in which “desires *de dicto* to do what is right, because it is right, seem to be reasonable as well as morally preferable” (Olson 2002 p. 92). One of the cases he presents he calls the “freedom of expression” case, in which a person is a strong supporter of freedom of expression because he believes it is morally right (Olson 2002 p. 93). The agent defends a Holocaust denier’s (“history revisionist’s”) right to speak and publish on views that the agent in question finds morally abhorrent. So he is not motivated by the *de re* desire to let this person speak their mind. Instead what he cares about is the rightness of freedom of speech, on Olson’s analysis:

This person is more likely to say that he cares about freedom of expression also *qua* its rightness...he may have strong *de re* desires that such opinions shouldn't be voiced, the 'history revisionist' should have his right to speak his mind because it is right that he should have that right...Rightness thus figures as a dominant part of the intentional object of his desire. It would be odd to say that such a *de dicto* desire would make this person morally flawed...it serves to make his position understandable and reasonable... (Olson 2002 p. 93).

The reason that the moral agent in this case is motivated to defend the history revisionist's right to speak is that he has the *de dicto* desire to do what is right, and in this case, defending free speech is right.

Another reason for thinking that *de dicto* desires to do the right thing may be a morally valuable part of an agent's moral psychology is that sometimes it may be that the *de dicto* desire to be moral is what motivates the agent to act in cases where the demands of morality are in conflict with other desires we have, especially the desire to act in our self-interest. Shafer-Landau writes, "...we can see the motive of duty as a special sort of standing desire, one that serves as a limiting condition on the formation of other desires" (Shafer-Landau 2000 p. 286).

Olson credits J. H. Sobel (in an unpublished manuscript) with making a similar point; ordinary moral agents, as opposed to ideal moral agents, are probably frequently motivated by *de dicto* desires to do the right thing. These desires, he says, are "somewhere in the back of [the agent's] head" and may "function as a kind of safety device when [the agent's] *de re* concerns are (momentarily) not sufficient to provide motivation to act" (Olson 2002 p. 92). Reversing Smith's charge of moral fetishism against the agent who is motivated by *de dicto* instead of *de re* desires, Olson suggests that it would be odd, and perhaps even morally undesirable for an agent to have no

concern (*de dicto* concern, that is) “about the moral rightness (wrongness) of her actions” (Olson 2002 p. 92).⁴⁴

Before concluding this defense of *de dicto* desires, it is important to consider a related point made by Shafer-Landau. Some internalists point to the phenomenology of moral motivation to suggest that *de dicto* desires are not the motivating force behind good actions in the good and strong willed agent. They claim “we can tell from simple introspection that many of our motivated moral actions are unaccompanied by such a desire” (Shafer-Landau 2000 p. 159, 285). Internalists claim that when we reflect on our own case we do not find that we are motivated by the motive of duty or the desire to be moral read *de dicto*.

Shafer-Landau rightly points out first, we can certainly be mistaken about what desires we have. We have no reason to think that we have perfect access to what our desires and motivations are. We shouldn’t think that all of the desires that cause us to act in particular ways have a “distinctive ‘feel’ at all” (Shafer-Landau 2000 p. 286). We have many “standing desires such as the desire to have a successful career, to have healthy parents, to avoid giving offense...” (Shafer-Landau 2000 p. 286). These are desires that we may not consciously reflect on very much at all. So it makes sense that we can have the standing desire to be moral without it jumping out on introspection (Shafer-Landau 2000 p. 286). Desires do not have to be the hot passions, instead “one may possess...desires without at every moment being aware of them” (Shafer-Landau

⁴⁴ If the externalist does embrace what Shafer-Landau calls the “motive of duty” and what Smith calls the *de dicto* desire to be moral as the favored explanation of the relationship between moral judgment and motivation, they still face the challenge of explaining why the presence of this desire is as widespread as it is. Shafer-Landau suggests that externalists have at least four options: innatism, socialization, self-creation and a hybrid account (Shafer-Landau 2003 p. 159).

2003 p. 158). In fact, if one takes desires to be functionally defined, an agent may have many desires that he or she is never aware she has.

The arguments given above are intended to make two main points. First, possessing and being motivated by *de dicto* desires to do the right thing is at least not morally objectionable, and at best, a morally praiseworthy characteristic. The example of Socrates is meant to show that there are plausible and widely respected accounts of moral motivation that do not necessarily involve *de re* concern for the objects of morality without being fetishistic. Olson's account of the free speech supporter also lends credibility to the notion that the *de dicto* desire to do the right thing is sometimes a virtue.

Second, Olson and Shafer-Landau's characterization of the *de dicto* desire to do the right thing as a background desire or limiting condition on the formation of other desires contributes to a picture of how a pluralistic externalist can incorporate non-fetishistic *de dicto* desires into their moral psychology. This defense of the *de dicto* desire to be moral fits with the most plausible version of externalism, pluralistic externalism.

3. Two Kinds of Externalism

Very broadly, motivation judgment externalism comes in two species. For both kinds of motivational judgment externalism, moral judgments are primarily expressions of beliefs, so they must be coupled with some kind of conative state in order to result in a motivation to act on them. What separates the first kind of externalism from the second is that in the first kind, the "conceptual content" of the "conative attitudes" "does

not involve moral concepts” (Svavarsdóttir 1999 p. 167). The conative state posited by this first kind of motivational judgment externalism is usually an emotion, like sympathy or compassion (Svavarsdóttir 1999 p. 167). I will refer to this type of externalism, after Zangwill, as no-moral content externalism. The second type of motivational judgment externalism, embraced by Svavarsdóttir and Zangwill, suggests that in order for moral judgment to motivate, it must be coupled with a conative state that does involve a moral concept or some moral content (Svavarsdóttir 1999; Zangwill 2003). For some externalists, this state is simply the desire to be moral (Svavarsdóttir 1999 p. 170). Let us call this second kind of externalism, moral content externalism (Zangwill 2003).

Brink’s externalism is a version of no-moral content externalism. He makes the important point that to a large extent people are motivated by their moral judgments because of both the beliefs and desires that the particular agent has and the moral theory that the agent endorses (Brink 1989 p. 49). He goes on to say that if the agent endorses a moral theory, for example, that emphasizes “the other-regarding character of many moral demands” then what may motivate this agent to act on his moral judgment is the experience of sympathy for his fellow man (Brink 1989 p. 49). Brink’s take on externalism seems to exclude the possibility of the motivating force of a desire with moral content, like a desire to do the right thing. Brink writes that “[e]xternalism claims that the motivational force and rationality of moral considerations depend upon factors which are external to the moral considerations themselves” (Brink 1989 p. 28).

Brink’s characterization of externalism seems too strong. Just because there is an external and contingent connection between the moral judgment and the motivation

to act on it, this doesn't have to mean that the motivation to act on moral judgments cannot contain any moral considerations (Zangwill 2003 p. 145).

Boyd's externalism also seems to be a kind of no-moral content externalism. Boyd agrees with Brink in that he claims that the moral theory that one endorses determines what will motivate an agent to act on their moral judgments (Boyd 1988 p. 186). Boyd ties motivation directly to the substantive moral theory that he defends: homeostatic consequentialism (Boyd 1988 p. 216). Boyd argues that moral goodness is tied to promoting a cluster of goods that "satisfy important human needs," such as physical well-being, health, love, companionship, cooperation, autonomy, intellectual work, etc. (Boyd 1988 p. 203). The intimate connection between making a moral judgment and having a motivation to act on it, can then be partially explained, according to Boyd, because human beings will "naturally prefer" actions and characters that promote these human goods (Boyd 1988 p. 215). Boyd thinks that sympathy, for example, provides a powerful explanation for why we care about other people achieving human goods (Boyd 1988 p. 215). It seems right to think that we have some desires to support actions and characters that promote important human goods. But Boyd's externalism also seems incomplete in that he does not address situations in which an agent lacks sympathy for the person positively affected by his actions, but does the right thing despite this. Nor does he address cases in which we have moral obligations that sympathy is inapplicable to, for example, moral obligations that we might have to ourselves or moral obligations that we might have towards the natural environment.

Zangwill's externalism serves as an example of the alternative kind of externalism, moral-content externalism. On this view moral motivation comes from a

distinct kind of desire, a desire with moral content. Zangwill argues that externalism should be framed in terms of “the representational content of propositional attitudes” (Zangwill 2003 p. 144). Those propositional attitudes could be either “cognitive or non-cognitive” (Zangwill 2003). If externalists use this language they can maintain that moral judgments are beliefs. In other words, externalists can claim that moral judgments are cognitive and have “realistic moral content... that is content that represent mind-independent moral states of affairs” (Zangwill 2003 p. 144).

At the same time, the externalist can say that “we also have non-cognitive states with realistic moral content-consider guilt, remorse, and the feeling of moral horror” (Zangwill 2003 p. 144). So Zangwill introduces two sets of propositional attitudes, those that make up our moral judgments and our motivating moral desires. When we are motivated to act on a moral judgment, it is because we also have an “independent desire” that also has moral content. The content of the desire that motivates us:

...represents a mind-independent moral state of affairs, where a moral state of affairs is the instantiation of a moral property by something. The motivating desire is the desire to do the morally preferable thing-or perhaps to do the right thing (Zangwill 2003 p. 144).⁴⁵

Zangwill rejects versions of externalism like Brink’s and Boyd’s, which claim that “desires with non-moral content,” like sympathy or benevolence are the sources of moral motivation, views. Zangwill has a clear explanation, consistent with the Humean theory of motivation about how moral actions are produced. Both the agent’s belief (her moral judgment) and her desire share the same or similar moral content (Zangwill 2003 p. 147).

⁴⁵ “Moral motivation, on this view, is a matter, first, of believing that some act of ours would have the property of being morally better than some alternative, plus second, there is an independent desire to do actions with that property” (Zangwill 2003 p. 144).

Zangwill's moral content externalism, what motivates us to act on our moral judgments is a distinctively moral desire. This desire is not moral in the sense that it is the only desire that is morally praiseworthy. Instead, it is distinctively moral in terms of its representational content. Zangwill calls non-moral desires that spring from empathy or compassion, "altruistic motivation[s]" (Zangwill 2003 p. 147).

... the distinction between moral and altruistic motivation, between our concern with duty and our concern with the welfare of others. Some actions are indeed motivated by sympathy and benevolence, but Kant was right to insist that "acting out of duty" or moral motivation is quite different. To describe actions that are motivated by duty or moral goodness as motivated by sympathy or benevolence is intuitively inaccurate. [...]The difference lies in the contents of the desires: one employs moral concepts, the other does not. The person motivated by duty has the thought "It is morally good that I do this thing," whereas the person motivated by altruism has the thought "So and so needs such and such." Altruistic acts are not done for the sake of duty, but for the sake of others. Acting out of altruism is one thing acting "out of duty" is another" (Zangwill 2003 p. 147).

Both kinds of desires could motivate the same action, for example, an agent gives to charity. It could be that the agent formed the belief with the content that giving to charity is their duty and had the desire to do their duty (a desire with moral content). At the same time the agent might feel compassion for the hungry children (a non-moral desire, according to Zangwill).

Despite Zangwill's recognition that both forms of motivation can be operating in an agent at the same time, and without giving one privileged status as being more morally praiseworthy, Zangwill sharply distinguishes between these two types of motivation: altruistic motivation is not a kind of moral motivation. It is clear from the above quoted passage that Zangwill takes Kant's distinction between the motive of duty and compassion and other inclinations that motivate us to be moral to be quite a significant one.

Recall that on Smith's view, the only kind of motivation that a virtuous agent could act on is the direct concern for the objects of morality (a *de re* desire). On Zangwill's view, that is called altruistic motivation. While Zangwill does not say that it is not virtuous to act on altruistic motivation, he does say that it is not moral motivation. Zangwill however, does not provide a reason for restricting the scope of moral motivation in this way. The pluralistic externalism that I defend can embrace both what Zangwill calls altruistic motivation and a host of desires with moral content all as forms of moral motivation.

3.1 Desiderata for Externalism

After getting a sense of both of the major strains of externalism, it is important to ask: What do we want from our view of the relationship between moral motivation and moral judgment? What features will the most plausible version of motivational judgment externalism have to have? I suggest three desiderata for the best theory of externalism.

First, externalism should be able capture the subtle complexity of moral motivation, making the first desideratum for externalism flexibility. Externalism should account for the appearance of a reliable connection between changes in judgment and changes in motivation and behavior on which Smith builds his internalism. When we change our minds about what the right thing to do is, we often also experience a change in motivation and a change in behavioral tendencies.

At the same time, externalism must be able to capture the darker side of our moral lives that is the reality of a range of forms of moral indifference and amoralism. This includes the variety in levels of motivation that obtain when competing different

agents, for example, you and I may share the same moral judgment that eating meat is wrong, but this moral judgment provides me with very strong motivation to abstain from eating it, while it provides you with only a weak motivation to abstain. But this variation can also be found within a single agent over time. There are also a variety of ways in which motivation to do the right thing can be destroyed under certain kinds of psychological stress and under pressure from competing desires (see Shafer-Landau 2003 p. 1460). Most controversially, externalism should be able to accommodate the existence of agents who systematically fail to be motivated by their moral judgments (Further examination of this variability will be undertaken in Chapter 3).

The second desideratum for externalism is the empirical constraint. Unlike many internalists, who generally argue for their position based on *a priori* and conceptual analysis of moral judgment, the externalist analysis of moral judgment presents itself as contingent fact about human moral psychology and moral discourse. Arguments for motivational judgment externalism are often presented as inferences to the best explanation. If externalists are to be consistent with their criticisms of internalism for focusing solely on conceptual *a priori* analysis, they must not neglect empirical constraints themselves. So the second characteristic externalism must have is it must be the one that fits best empirical evidence, taken from psychology, cognitive science and neuroscience, about the relationship between beliefs, and motivation.

The best version of externalism should not include a view or make assumptions about what make a good agent good, or morally praiseworthy, this is the neutrality constraint. This question should be left as open as possible, because we want a theory of moral psychology and motivation that can fit with a variety of normative theories

about what makes someone virtuous, etc. First order normative theory may side with Kant that an action is only morally praiseworthy if it is produced by the good will, and done out of respect for the moral law or it may side with Mill and other consequentialists, that an action's moral worth is the same irrespective of the agent's motivation for performing it. It is better if externalism remains neutral on this issue because these are in principle two distinct questions: how does morality motivate us and how should morality motivate us? The third desiderata is thus normative neutrality.

3.2 Pluralistic Externalism

The version of externalism that best fits these conditions, the flexibility condition, the empirical evidence condition, and the neutrality condition is a pluralistic or hybrid externalism. It is pluralistic in that it embraces a range of kinds motivations. It does not assume that there is one dominant or uniform type of interaction between these mental states. Being motivated to act on one's moral judgments is a contingent matter. Pluralistic externalism provides sufficient conditions for being motivated to act on one's moral judgment; it does not provide necessary conditions for being motivated to act on one's moral judgment. The pluralistic view can embrace both Zangwill's and Svavarsdóttir's style of externalism, as well as Brink's and Boyd's. For example, on the pluralistic view, sometimes an agent may be motivated by the *de dicto* desire to do the right thing. Other times, it may be that the agent's sympathy for another person causes them to act on their moral judgment.

In this way, pluralistic externalism is close to Mill's account of the motivations that people have to be moral (Mill 1863). The external sanctions include fear of punishment

and hope for reward, from other people and God, fear of social ostracism, and the “sympathy and affection” that we feel for our fellow human beings (Mill 1863). Mill urges that feelings of sympathy and “unity with our fellow creatures” should be inculcated through education (Mill 1863). The internal motivation to be moral is also a mental state, but it is not sympathy. Instead it is:

...a feeling in our own mind; a pain, more or less intense, attendant on violation of duty, which in properly cultivated moral natures rises, in the more serious cases, shrinking from it as an impossibility (Mill 1863 p. 28).

This kind of motivation is contingent; “[u]ndoubtedly this sanction has no binding efficacy on those who do not possess the feelings it appeals to (Mill 1863 p. 29).

On the pluralistic account an agent may be motivated by the *de dicto* desire to do the right thing. This is a standing or background desire. An agent may also have direct desires or concerns for other morally relevant objects like, being a good friend, the happiness of other people, or global justice. These desires or concerns may also provide them with motivation to act on their moral judgments. There may also be a relationship between making a moral judgment and then acquiring a desire to promote the things that one judges to be morally right.

Pluralistic externalism best fulfills the three desiderata for externalism. The first part of the flexibility condition is meeting a standard close to Smith’s reliable connection standard. It does surprise us when other people espouse a moral view and then claim they have no motivation to act on that view. This is easily explained on pluralistic externalism. There are many different ways people can be motivated to act on their moral judgments, so they often do have some motivation to act on them, whether they are motivated by the desire to do the right thing, empathy or some other concern. The

other side of the flexibility condition is that externalism must capture the observation of variable motivation in one agent at different times, in different agents, the way motivation can be snuffed out and the existence of the morally indifferent. Pluralistic externalism can also accommodate these complications, as a form of externalism, the theory says that it is perfectly possible for an agent to make a moral judgment and fail to be motivated by it.

Pluralistic externalism has the best prospects of meeting the empirical evidence condition. We know that people who have diminished empathy and other affective limitations have trouble acting on their moral judgments, even when there is some evidence that they seem to make the same moral judgments that the rest of us make. We also know that empathy, sympathy, and compassion can sometimes lead people to act against what they judge to be morally required of them. Most importantly, pluralistic externalism is best equipped to meet the empirical evidence condition because it embraces its status as an *a posteriori* view that is open to amendment.

Pluralistic externalism best meets the neutrality condition because it does not build into its content, the way that Smith or Zangwill do any judgments about what a uniquely “moral” in the sense of morally good or praiseworthy motivation is. Pluralistic externalism aims to be a purely descriptive theory and can remain compatible with a variety of normative ethical theories about motivation.

For example, on some interpretations of Kant, acting on the good will is the only type of moral motivation that can confer morally praiseworthiness on an action. Kant explicitly says that in reality it is very difficult to know whether there ever are any truly morally praiseworthy actions.

In fact there is absolutely no possibility by means of experience to make out with complete certainty a single case in which the maxim of an action that may in other respects conform to duty has rested solely on moral grounds and on the representation of one's duty... We like to flatter ourselves with the false claim to a more noble motive; but in fact we can never, even by the strictest examination, completely plumb the depths of the secret incentives of our actions (Kant 1785/2012 Sect. 2, Par. 407).

Versions of externalism that stipulate up front what kind of motivation the good agent has, or what kind of motivation is moral motivation disallow for this kind of uncertainty.

What I take to be a strength of pluralistic moral judgment externalism, some would argue is a weakness. A possible objection to pluralistic externalism may be that it allows for too many motivations, even selfish ones to count as 'moral' motivation. The question arises, does pluralistic externalism need to be able to distinguish desires that motivate a person to do the right thing from desires that motivate a person to do what happens to be the right thing to do?

It seems that being motivated to act on one's moral judgments is not the same thing as just having a moral judgment and also having a motivation that leads one to do the same action that the moral judgment would prescribe. For example, someone might believe that the right thing to do is to help an old lady cross the street. He lacks any empathy for the old lady, does not have a desire to ease human suffering (or any like desire), and does not have a desire to do the right thing, qua the right thing. Instead, he has a desire for a reward. He believes that if he helps the old lady cross the street she is likely to give him a reward. So this agent is certainly motivated to help the old lady cross the street. He also believes that this is the right thing to do. But he does not seem to have moral motivation; he is not motivated to act *on* his moral judgment. He is exactly

the sort of person that internalists have to explain away because he is exhibiting a form of moral indifference or amoralism.

An explanation of the difference between moral motivation and this agent's motivation cannot merely rely on the counter-factual claim that in this case, if the agent did not make the moral judgment 'helping the old lady cross the street is the right thing to do,' he would still be motivated to help the old lady cross the street. This explanation does not help. This is because the same thing could be said about motivations based on compassion, or a direct desire to ease the suffering of others, and those are motivations that pluralistic externalism counts as moral motivations. We could ask, if the compassionate individual didn't make the moral judgment, 'helping the old lady cross the street is the right thing to do,' would they still have helped the old lady cross the street? The answer to this question may very well be, yes they would have. This does not seem like the appropriate grounds on which to say that they weren't acting on a "moral motivation." If I were to adopt the view that the only motivations that count as distinctly moral are those that are desires with moral content, this puzzle could be avoided. This is not a puzzle that I can respond to here, except to repeat that this is an issue for first order normative theory.⁴⁶

Zangwill raises a related objection to his own view:

Objection (b): If actions are always the result of desires in addition to beliefs, then they are motivated by selfish considerations. But moral and selfish motivations are diametrically opposed. So the motivating force of moral judgments cannot be external to them (Zangwill 2003 p. 145).

Zangwill responds to this objection by pointing out that externalism does not entail psychological egoism, the view that people are only ever, at bottom motivated by self-

⁴⁶ However, I am not arguing that metaethics and moral psychology can take a completely Archimedean perspective on the rest of ethics.

interest. Of course, I agree with Zangwill on that point. Psychological egoism is a notoriously difficult position to defend, as it makes the sweeping claim that all of our desires are self-interested, rather than being directed at the well-being of others, for example.

4. Conclusion

In Chapter 1 I introduced a set of five claims that moral realists make. Defending naturalist moral realism against the internalism objection and arguing for motivational judgment externalism helps to secure for the moral realist claim (3), that is, psychological cognitivism, the claim that moral judgments are beliefs. In this chapter moral realism was defended against the internalism objection in several ways: first, by undermining the considerations in favor of internalism. Second, the chapter posed an independent argument against internalism. And third, I presented a plausible version of externalism, pluralistic externalism that can adequately meet the challenges that internalism raises. In the following chapter I will present another independent argument against internalism, the argument from moral indifference.

Chapter 3: Moral Indifference and Amoralism

The morally indifferent are those who whether globally, locally, permanently or impermanently do not “car[e] very much about the demands of morality” (Zangwill 2008 p. 101).⁴⁷ Characters exhibiting a wide range of forms of moral indifference pose problems both for internalism about moral motivation, and non-cognitivism more generally.⁴⁸ The morally indifferent play a significant role in arguments against internalism, because they seem to be able to make the same moral judgments that morally deferential people do, without having any corresponding motivation to act on those judgments. Thus at their strongest, the morally indifferent present a central counter-example to the internalist’s claim that there is a necessary, essential, or inherent connection between moral judgment and motivation.

The morally indifferent also pose a problem for some non-cognitivist theories of moral thought. This is because non-cognitivists generally characterize moral judgments as conative or affective states, which are usually taken to be inherently motivational states. If the morally indifferent can make a moral judgment without having the corresponding motivation, then this casts doubt on the notion that moral judgments are conative states.

Cognitivists who are also externalists can give a more straightforward explanation of the morally indifferent than anyone else can. On externalism, the moral judgment itself does not motivate the agent to act; instead, it is some appropriately

⁴⁷ Stefan Baumrin, in “Moral Blindness,” argues that moral philosophers have by and large failed to put sufficient effort into explaining cases of gross failure to act morally or make the right moral condemnation of those who act in a grossly immoral fashion (Baumrin 1986).

⁴⁸ In the previous chapter I focused on the weak appraiser form of motivational judgment internalism. So in this chapter I use the term ‘internalism’ to refer to the ‘weak appraiser form of motivational judgment.’

related conative state, like a desire, that motivates. The relationship between the moral judgment and the motivation is a contingent one. On this view, moral indifference occurs anytime the appropriate conative state is missing or fails to connect to the moral judgment in the relevant way.⁴⁹

In the previous chapter, I suggested that my proposed version of externalism, pluralistic externalism, has advantages over internalism, as well as over other more narrowly conceived forms of externalism. The fact that we are often motivated to act on our moral judgments is a purely contingent matter. The view is pluralistic in at least two ways. It does not assume that there is one dominant or uniform type of interaction between moral judgments and moral motivation. Nor does it stipulate that only certain kinds of conative states motivate people to act on their moral judgments. For example, on the pluralistic view, sometimes an agent may be motivated by the desire to do the right thing. Other times, it may be that the agent's sympathy for another person causes them to act on their moral judgment. In other words, pluralistic externalism provides sufficient conditions for being motivated to act on one's moral judgment but it does not provide necessary conditions for being motivated to act on one's moral judgment.

Because on this view agents are motivated to act on their moral judgments by a wide variety of sources, there are also a great many ways this motivational system may fail to function. Without a necessary, essential, or inherent connection between moral judgment and motivation, failures of motivation are much easier to explain. And if we grant pluralistic externalism, the range of possible explanations for motivational failures is wide open. For example, an agent may lack the desire to do the right thing in general

⁴⁹ The externalist who is also moral realist also has the unique advantage when it comes to explaining amorality, in that they can give a better account of the nature of the disagreement between the moralist and the amoralist (Brink 1989 p. 85).

or in a particular case, or an agent might not desire something more specific, like the well-being of her friend in a particular situation. Some agents are highly motivated by their moral judgments and others are only motivated a little or not at all because they have different psychologies.

In this chapter I will first present a brief taxonomy of some of the major forms of moral indifference and explain the role that they each play in arguments for externalism and against internalism. In the following section I will move from the most extreme case of moral indifference, amorality, to the most mundane, everyday cases of occasional amorality and diversity in levels of moral motivation. I also include a brief section on depression and listlessness. Although I discuss the amoralist, depression, and listlessness, along with others, I stress that moral indifference should not be understood as an exotic phenomenon found only in psychopathic serial killers or suicidally depressed. Instead, it is commonplace. I suggest that moral indifference might actually be the majority of our moral experience, rather than the exceptional case.

Second, I will explain how internalists have defended their view against objections based on moral indifference and amorality. The basic internalist strategy is to come up with an alternative explanation of the morally indifferent. The most popular strategy is to suggest that the externalist presents a distorted picture of what the morally indifferent are doing when they make a moral judgment (construed either as a mental act or a verbal act). I will assess the internalist responses, finding them lacking on grounds specific to each argument, as well as on the general grounds that they fail to meet the standard of giving independent support—an independent reason to believe—their reinterpretation of the cases of moral indifference. Finally, I revisit pluralistic

externalism and provide some reasons why it has a straightforward explanation of moral indifference.

1. A Brief Taxonomy of Moral Indifference

1.1 Amoralism

Because for the amoralist, the necessary connection between moral judgment and motivation has been severed, externalists cite the amoralist as evidence against internalism. The amoralist is not dissuaded from performing an action by the fact that he judges that the action is morally wrong. An amoralist is different from someone who just does the wrong thing, occasionally, or even consistently, someone we might call a bad person. Such a person might simply have very strong counter-veiling motivations that always or often outweigh the motivation he has to act on his moral judgment. A bad person might have a host of false beliefs about, either what their circumstances are, or about what the consequences of their actions will be, or about what the right thing to do actually is.

The amoralist, on the other hand is not motivated by his moral judgments at all. Amoralists do not necessarily, although they might, engage in a lifetime of immoral behavior. For instance, we can imagine, or may perhaps know someone, who behaves morally as long as doing the right thing is consistent with their prudential interests. However, when morality and prudence diverge, this person does the prudential action. All along this agent hasn't been motivated to act on his moral judgments; he has been motivated only to act on his prudential judgments, which happen to coincide with morality. It seems that not only can we imagine an amoralist, but we also have

evidence to suggest that such people exist. How the behavior and utterances of amoralists should be interpreted plays a role in a wide range of metaethical debates.⁵⁰

Peter Railton's case of Roger, the "sensible cad," is a good example of an amoralist, who can make moral judgments in a reliable and seemingly normative way but is not motivated by them.⁵¹ Roger is a charming and attractive man who works in an office. He sleeps with many of the women in his workplace and then leaves them. When he is criticized by his colleagues, Roger explains his behavior to his co-workers in the following way: "Look, we're all grownups—they know what they're doing. I never force myself on anyone—that just turns me off" (Railton 2009 p. 170). The boss in this office is Fred, who has been "putting a lot of pressure on an attractive new secretary, Marisha, to go out with him, and she's begun to yield," something he has done before (Railton 2009 p. 170). Roger joins in the following water-cooler conversation about the relationship developing between Fred and Marisha:

'...Marisha wouldn't have anything to do with Fred if he weren't her boss and putting all kinds of pressure on her. When someone with direct power over you puts moves on you, it's got to be harassment.' Instantly Roger is challenged and hominen...Roger rejoins. 'I never have any personal power or official authority over the women I go out with. And I don't want any. What would be the fun in that? Like I say, coercion is a turn-off. Except maybe for jerks like Fred. What Fred's doing is harassment because he's got the power to hire or fire her,' he continues. 'Yes isn't yes in a case like that. But don't think that you'll do Marisha any favor by reporting it. Believe me, I know this place. You'll just get her in trouble.' A co-worker is put off by the cad's unfeeling tone. 'So you're happy with just letting this go on, knowing that it's harassment?' 'I'm neither happy nor unhappy,' he replies. 'It's nothing to me. For losers like Fred, harassment is the only way to get to first base. But if Marisha's as smart as I think she is, she knows this, and she'll play him for all he's worth' (Railton 2009 p. 170).

⁵⁰ Non-cognitvists of all stripes often cite the moral psychology of amoralists as evidence for the close connection between having the right sentiments and doing the right thing. See (Prinz 2007; Nichols 2002).

⁵¹ Also see Zangwill's mercenary, and Svavarsdóttir's Patrick, the moral cynic (Zangwill 2008 p. 106; Svavarsdóttir 1999).

As Railton presents the case, Roger is not misusing moral language. Roger is sincere in saying what he thinks and feels. He seems to understand the term 'sexual harassment' well and why it is wrong. Roger is able to speak in an authoritative way about moral issues. He can make arguments and has facility with moral concepts. Roger does not sexually harass women, but this is not because of moral concerns, he just does not enjoy it.

Roger provides a useful characterization of an amoralist because he does not behave in an overtly immoral manner and he seems to have a clear grasp of moral concepts, and is even at times, morally insightful. Given this characterization we have no reason not to think that Roger makes moral judgments. Despite this, Roger seems to fail to be motivated by his moral judgments. Whether or not people like Roger exist in reality, we can certainly conceive of such people. Thus Roger and others like him are counter-examples to the view that there is a necessary connection between moral judgments and motivation, internalism.

1.2 The Occasional Amoralist

Amoralists conjure up images of people who are evil and do very bad things. But an amoralist does not have to be someone who always fails to be motivated by their moral judgments or someone who fails to be motivated by their moral judgments in dramatic or extreme moral situations. Individual instances of amorality may also exist in a single person. In order to bolster the argument against internalism, as Shafer-Landau writes, "[w]e need defend only the conceptual possibility of an agent who on a single occasion fails to be motivated by a moral judgment that he endorses...the moral

indifference we assign such a person needn't be systemic" (Shafer-Landau 2003 p. 145). Shafer-Landau canvasses several cases to help us imagine under what circumstances this could occur.⁵²

Some moral demands, may be so strenuous and difficult to fulfill that the agent loses all of their motivation to comply with them, yet continues to judge that he or she is behaving immorally by failing to comply: "[a] sincere prediction of futility may be sufficient to entirely sap one's motivation for carrying through...such a prediction does not appear to entail revocation of the value judgment" (Shafer-Landau 2003 p. 149). This is certainly the case when it comes to the accomplishment of non-moral goals. We may be initially motivated to work to achieve a goal, but as it becomes apparent that there is a high probability that we will not succeed, our motivation begins to wane and sometimes disappears entirely. Shafer-Landau suggests we have no reason to think that moral goals would be any different. Similarly, there is the cynical agent who makes the moral judgment that she ought to donate to famine relief, for example. But she also believes that since no one else is doing the same thing, her own efforts are pointless and thus loses all motivation to give to famine relief (Björnsson 2002 p. 339).

Second, in cases where prudence is at odds with the demands of morality, that is doing the right thing involves great personal sacrifice, an agent may lose all motivation to comply with his or her moral judgment (Shafer-Landau 2003 p. 149). For example, imagine a soldier who is ordered to go to the front lines of a battle where he is very likely to be killed. His fear of death may snuff out his motivation to go to the front lines and instead he might desert. He may still judge himself a coward. This judgment reveals

⁵² Shafer-Landau takes it to be a virtue of his cases that they are "neutral as between Humean and anti-Humean assumptions about moral judgment" (Shafer-Landau 2003 p. 149).

that he continues to hold that the morally right thing to do, his duty, is to join his fellow soldiers on the front lines (Shafer-Landau 2003 p. 150).

Third is the freshman philosophy student who is fully persuaded by ethical relativism.

Imagine an introductory philosophy student who has become convinced of the truth of a sort of ethical relativism. She believes that the ultimate moral standard comprises the fundamental mores of the society in which an action is performed. Armed with this view of morality, she issues certain moral judgments that she takes to be correct. But she is alienated from her society...she finds much of the prevailing cultural code amendable, she rejects a strand. She is voicing what she takes to be the moral truth, yet is unmoved (Shafer-Landau 2003 p. 150).

Copp's case of Alice is similar in that the overall moral system that she believes in yields a particular moral judgment that she fails to be motivated by (Copp 1995). Alice is a divine command theorist, she believes that right and wrong are "determined by the commands of God"(Copp 1995 p. 190-1). Part of her conception of God is that of a "vengeful ruler" who endorses the dictum "an eye for an eye," as a result of this conception of God:

Alice believes that capital punishment is obligatory in cases of murder and she believes she has an obligation to support capital punishment. But she is deeply compassionate, and she is quite out of sympathy with what she takes to be God's vengefulness. Because of her compassion she is not motivated in the least to support capital punishment. She is in fact active in opposing it, even though she believes she is morally forbidden to do so (Copp 1995 p. 190-91).

Shafer-Landau's fourth example is the case of the moral outsider. This person used to make moral judgments like the rest of us and be accordingly motivated. "But something awful happens to this person, causing her to reassess her attachment to morality...she no longer feels compelled to 'play that game'" (Shafer-Landau 2003 p. 151). Although this person continues to make the same moral judgments that she has been making her whole life, she is now detached or alienated from morality in the sense

that she no longer “gives any sort of priority to moral demands” (Shafer-Landau 2003 p. 151).⁵³

1.3 Diversity in Moral Motivation

You and I may both judge that it is morally wrong to eat meat. This might give you a strong motivation to abstain from it, while it gives me only a weak motivation to abstain from eating it. There is wide variation between individuals in the degree of motivation that they have even while making the same moral judgment.

On a more general level, moral concerns might motivate me to a greater extent across my lifespan, than they motivate someone else, even though we share most core moral beliefs (Zangwill 2008 p. 101). In other words, morality might play a greater motivational role in one person’s life than another’s. This is not the claim that different individuals have competing desires and will therefore not always end up acting on their moral motivations to the same extent or in the same situations. That is very likely true, but probably not enough to explain the phenomenon. Instead, the objection is that the force of the motivation that the moral judgment provides seems to be stronger in some people than others.

Moral judgments can also have a varying degree of motivational force in one individual across time (Zangwill 2008 p. 101). Zangwill’s mundane example is that when he first wakes up in the morning, before having his coffee, he holds all of the

⁵³ Russ Shafer-Landau: “externalists claim that the connection between holding a moral judgment and being motivated is a contingent one...contingent on a person’s psychological make-up and on the perceived content of moral demands” (Shafer-Landau 2003 p. 151). The case of futility, the conflict with prudence and the freshman relativist are cases in which the “perceived content of moral demands” “undermines” motivation to act on moral judgments. Cases of depression and listlessness and of the person who has lost her “attachment to morality” are cases in which “aberrations in psychological make-up...cut the connection between moral judgment and motivation” (Shafer-Landau 2003 p. 151).

same moral beliefs that he did the day before, but they have much less motivating force than they do at any other time of the day. In the simplest case, in a life full of moral judgments that lead to moral action (in whatever way that is cashed out), an agent may on a *single occasion* fail to be motivated by their moral judgment (Bedke 2009 p. 191).

Other times there is a delay between the making a moral judgment (especially when that moral judgment is a reversal of a previously held moral judgment) and when we are motivated to act on that moral judgment, or when it informs our behavior (Sadler 2003 p. 72). For example, if for a long time I believed that eating meat is morally permissible, and was subsequently convinced of the immorality of eating meat, it may take time and some effort on my part to work up the motivation to actually stop consuming meat. At times, it requires considerable and “deliberate effort” to motivate ourselves, to rev ourselves up, to act on our considered moral judgments (Sadler 2003 p. 72). If moral judgments necessarily, essentially, or inherently motivate, as internalists claim, this delay needs to be explained. On the internalist view we would expect that a change in moral judgment would be “automatically” accompanied by a change in motivation, meaning that it would happen immediately and without effort (Sadler 2003 p. 72).

I am grouping these phenomena of variability together with moral indifference because of the way they similarly challenge the internalist claim that there is a necessary or inherent connection between moral judgment and motivation. Given the phenomenon of variability in moral motivation, the internalist has the further burden of explaining why it is that different people are motivated to varying degrees by the same moral judgments. This is not to say that it will be impossible for the internalist to come

up with some explanation of these cases. They just have more work to do. If the internalist is a non-cognitivist, and claims that moral judgments are in fact, just attitudes of some kind, which are inherently motivational, then “why should the motivational aspect of attitudes vary?”(Zangwill 2008).

Another often overlooked form of diversity in moral motivation that supports externalism has to do with the wide range of kinds of moral judgments that we make. Many of our moral judgments involve speculating about hypothetical cases, or commenting on the rightness or wrongness of the actions of other people, either people around us, our politicians or world leaders, fictional characters, or the actions of historical figures (Sinnott-Armstrong 2010 p. 66). We also make judgments about states of affairs being (or having been) good or bad and types of actions, thoughts and characters being right wrong, good, bad, praiseworthy, or blameworthy etc. And finally, we make judgments about our own conduct, in many temporally diverse ways. The judgments we make that pertain to ourselves range widely from judging how we have behaved in the past, deciding how to behave in the very long term, as in reflective life-planning and considering which values to prioritize, to concrete immediate decision-making about whether or not to tell a lie to get out of *this* awkward situation (Sinnott-Armstrong 2010 p. 66) .

In many of these cases it seems implausible to say that the moral judgment necessarily, essentially, or inherently motivates. I suggest that many of the moral judgments that we make involve this kind of moral indifference. For example, the necessary connection with motivation seems strained when reflecting on the actions of other people that occurred in the distant past. How would reflecting on the injustices

perpetrated by a particular feudal lord against his serfs in medieval Europe necessarily motivate me, and to do what? Internalists might say that it would motivate me to prevent those conditions to ever exist again, for example. But that surely involves another, second, or third moral judgment, namely that the same actions would be wrong if, in some roughly analogous situation in the present day, and further that if such a situation were to develop I would be obligated to do something about it. Even if internalists could make a convincing case that all of these moral judgments necessarily, essentially, or inherently motivate, it seems unlikely that they would they all have equal motivational force.

An initial response that internalist might give to these cases of motivational variability and indifference is to limit the scope of internalism by restricting the kinds of moral judgments that the theory applies to. For example, they may claim that internalism only applies to “first-person present-tense (in situ) moral thoughts that seem true”(Sinnott-Armstrong 2010 p. 67). While this weak form of internalism limits the number of cases that the internalist has to be able to offer a plausible interpretation of, it does seem like an ad hoc limitation of the view (Sinnott-Armstrong 2010 p. 67).

The other important point to notice about this limitation is that it makes internalism inapplicable to a very wide swath of our moral judgments. Assuming that many other people are like me in this way, in the course of an average day, I may make very few first person, present-tense (in situ) moral judgments, but I make many judgments about other people and states of affairs. Embracing this restricted form of internalism would make it very narrow in scope. This is especially problematic if the truth of internalism is being used in an argument against cognitivism and thus against

realism. Even if we were to grant the truth of internalism in a very restricted set of cases, this would provide little evidence for a claim about the nature of moral judgments in general.

The pluralistic externalist position, which I defended in the previous chapter, can give a straightforward account of the variability of moral motivation in its various forms. How much or whether a moral judgment motivates an agent has to do with “distinct desire[s]” and other kinds of conative states, which have different levels of strength (Zangwill p. 103).

Another widely discussed example of agents who lose their motivation to act on their moral judgments, but continue to make them are people who are depressed or listless.

1.4 Depression and Listlessness

Moral indifference is also sometimes thought to be the result of depression and listlessness (Dancy 1993 p. 5). It is more difficult to make the case that depression and listlessness are counter-examples to internalism, than other forms of moral indifference, because they are in many cases pathological states that are not well understood (Zangwill 2008 p. 113).

However, I include this brief discussion of depression and listlessness because first, it expands the range of phenomena that the internalist has a harder time explaining than the externalist. There is a difference between being clinically depressed and being depressed. While clinical depression has a complex psychiatric etiology and neurochemistry, everyday depression is something that most of us are familiar with and

can make a host of folk psychological claims about.

Second, it might be that depression (clinical or otherwise) is related to a wholesale failure of the depressed person's motivational system and that the depressed are not motivated by their prudential judgments either. They are not motivated to take a shower or go to therapy, even though they judge that these things will make them feel better. They may not even be motivated to attend to their basic needs (the activities of daily living, in the language of psychiatry). A large scale motivational failure does not rule out the depressed or listless as counter-examples to internalism. If it is possible to lose motivation to do much of anything, including acting on one's moral judgments, and yet continue to be able to make moral judgments, then presumably there is a disconnect between the two, as the externalist contends. In response to these cases, many internalists restrict their view, to accommodate such failures of motivation, and included a practical rationality restriction on who their version of internalism applies to, such restrictions will be discussed in internalist responses section.

Alfred R. Mele uses the character of the listless person to argue against a form of internalism (Mele 1996 p. 731). The listless agent that Mele describes is "clinically depressed" yet retains some of her moral beliefs, "while being utterly devoid of motivation to act accordingly" (Mele 1996 p. 733). This agent, whom he calls Eve, can hold a moral belief that she ought to help her Uncle, and yet, due to listlessness caused by the death of her husband and family, fail to be motivated by that belief. According to Mele, to deny that Eve can both hold this moral belief and fail to be motivated by it is just to assume internalism (Mele 1996 p. 736).

As discussed in the previous chapter, one possible source of the motivation to act

one one's moral judgments could be the "desire to do whatever is morally required" (Mele 1996 p. 733). Mele imagines two different planets; on one planet, Planet X, almost everyone possesses the very strong desire to do whatever is morally required. Very rarely, the inhabitants of this planet succumb to listlessness and their desire to do the right thing wanes. In those cases they fail to be motivated to act on their moral judgments. On Planet Y on the other hand, everyone is "...uniformly equipped, genetically, with a generic desire of the kind just described... the desire is so strong that it is infeasible by any competing motivation"(Mele 1996 p. 738). On planet Y "MR [morally required] beliefs are uniformly efficacious" (Mele 1996 p. 738). The difference between the two planets is that on planet Y, unlike on planet X, everyone always has the desire to do the right thing, and thus never fails to be motivated by their moral judgments.

Mele's point in making this comparison is that even if we earthlings are like the inhabitants of planet Y, rather than planet X, internalism could still be false. Even on planet Y, motivation is only contingently, not necessarily tied to having a moral belief/making a moral judgment. In other words, on planet Y, things could have been otherwise. Internalism requires that they be necessarily connected.

2. Internalist Responses to Indifference

The conceivability or existence of people who can make moral judgments but are not motivated to act on them is a counter-example to the theory of that there is a necessary, essential, or inherent connection between moral judgment and motivation. In the previous section I offered a range of ways in which someone might fail to be

motivated by their moral judgments. Internalist replies have focused primarily on responding to the case of the amoralist who consistently fails to be motivated by his or her moral judgments. At the end of the responses section, after I have argued that these replies fail in various ways, I will explain why these failures are particularly apparent, when applied to cases of variable motivation, occasional amorality, listlessness and depression.

The internalist has four possible responses to the amorality counter-example.

(1) She can restrict the cases in which internalism applies so that the amoralist falls outside of the realm of cases that internalism needs to explain, by excluding people who are not “good,” or “strong willed,” or who are “irrational,” or “practically irrational” (Lockie 1998 p. 21; Smith 1995).

(2) She can deny that the amoralist makes moral judgments at all. This is by far the most popular strategy for responding to the externalist objection – re-describing what the amoralist is actually doing. Internalists may accept that the amoralist does lack moral motivation, but they argue that we are mistaken in saying that the amoralist is really making a moral judgment. For example, the claim originally made by R.M. Hare, is that the amoralist is not making the moral judgment at all; he is simply stating the fact that *other* people judge, or it is conventionally accepted, that the act in question is morally wrong (Hare 1952 p. 125).

(3) The internalist may concede that while individual instances of moral indifference or amorality are possible, it is impossible to conceive of a community of amorality.

(Bedke 2009). This is sometimes taken to show that the inverted comma interpretation of the amoralists moral judgments is the correct one (Lenman 1999; Bedke 2009).

(4) The internalist can deny that the amoralist lacks motivation; this is usually done by re-construing what motivation amounts to, or by arguing that the amoralist is in fact motivated, but very little, by his moral judgment (Björnsson 2002).

The second, third, and fourth responses attempt to diffuse the amoralist counter-example by offering an alternative interpretation of his or her behavior, utterances, and psychological states. In order to be successful these responses have to be able to show that there are independent grounds for accepting them, besides an effort to defend internalism (Lockie 1998 p. 22). It is on this ground that internalist responses to cases of amoralism overwhelmingly fail. For example, there have to be good independent reasons for saying that in a particular case an agent does not make a moral judgment, other than the fact that this is the only way to make the evidence fit with internalism. Most of the time it makes sense to take what people say about how they feel, what they believe, and what they are experiencing as good *prima facie* evidence about how they feel, what they believe, and what they are experiencing. These internalist responses have to rely on exceptions to this principle.

Externalists are not the only ones who point out that there have to be independent reasons for favoring one interpretation of the amoralist's moral judgments over another; this point is also made by Michael Smith. He thinks that he *has* provided the grounds for favoring his internalist interpretation of the amoralist's moral judgments (See Chapter 2 Section 2.3) (Sadler 2003 p. 68-9). However, the independent reasons

he gives for favoring his interpretation of the amoralist have been undermined by arguments I made in Chapter 2 having to do with *de dicto* versus *de re* desires to do what one is morally obligated to do and the charge of moral fetishism (See Chapter 2 Section 2.3).

It is difficult to establish who bears the burden of proof in the debate. There is at least one methodological principle taken from Svavarsdóttir that suggests that the internalists bear it (Svavarsdóttir 1999 p. 178-179). She argues that in any field when two reasonable people disagree about which hypotheses is the correct explanation of some observable phenomenon (in this case the behavior and utterances of the amoralist or the morally indifferent) the burden of proof is on the person who insists on the more “restrictive explanation” of the phenomenon in question (Svavarsdóttir 1999 p. 178-179).

In this case the externalist says that what might be going on in cases of moral indifference is that these agents are making moral judgments, but failing to be motivated by them. In defending the necessary connection between moral judgment and motivation, internalists say that their opponents’ explanation is impossible. The externalist interpretation is immediately ruled out because it is inconsistent with the internalist proposed concept of a moral judgment. Thus the internalist gives the more restrictive interpretation of indifference and amoralism, they eliminate the externalist understanding of the situation in virtue of their conception of moral judgment.

2.1 Restricting the Scope of Internalism

Internalism often serves as a crucial premise in one type of argument for the view that moral judgments are non-cognitive or conative states. If moral judgments necessarily, essentially, or inherently motivate an agent to act on them, then they cannot be beliefs, the argument claims. Since beliefs are not the type of things that (by themselves) can motivate an agent to act.⁵⁴ But as internalists restrict the scope of agents that the theory applies to, the role that it can play in metaethical disputes about cognitivism diminishes (Björklund et al. 2012 p. 126). The more restricted the form of internalism, the smaller role it can play in arguments of this kind. For example, on a restricted form of internalism, moral judgments may be necessarily, essentially, or inherently, tied to motivation only in agents that are morally good, practically rational, and mentally healthy (lacking depression, etc.). Further, the theory only applies to moral judgments that are in the first person, and applied only to actions that I am about to engage in or refrain from doing. This characterization of internalism severely restricts the class of agents and judgments that the theory applies to. For this reason such a restricted form of internalism would seem to give us more of a reason to think that moral judgments and motivation are two separate kinds of mental states that are complexly related to each other, rather than thinking that moral judgments are expressions of non-cognitive states like approval or disapproval.

One of the most common restrictions on internalism is a rationality constraint. The rationality constraint is a convenient way for internalist to respond to amoralism or moral indifference –these agents are irrational, and thus beyond the scope of internalism (Smith 1995; Korsgaard 1996; Garrard and McNaughton 1998; Dancy

⁵⁴ I am not disputing the claim that beliefs cannot by themselves produce action (part of the Humean theory of action).

1993). To paraphrase, Walter Sinnott-Armstrong, the morally indifferent go from being impossible, to irrational (Sinnott-Armstrong 2010 p. 69). There are several different versions of this view, but the basic claim is that it is a requirement of practical rationality that one is either motivated by one's moral judgments, or that one acts on one's moral judgments.

It is beyond the scope of this chapter to assess the individual merits of each of these views. And much of the plausibility of any one of the views may come down to what these internalists include in their theories of practical rationality. They all bear the burden of making a convincing case that the morally indifferent are irrational. Not only is this a more complicated account than the one that the externalist can give, there is little evidence to think that one has to be irrational in order to fail to be moved by one's moral judgments (Zangwill 2008 p. 115; Foot 1958, 1972). To claim that the morally indifferent are irrational the internalist has to bite a bullet. If moral indifference is as widespread and common as I am claiming it is, then we have to accept that many of us, much of the time, are irrational. We have to discount all appearances to the contrary, even in more extreme cases, like the case of Roger. Second, it is unclear how much of a help theories of practical rationality actually are in explaining these cases of moral indifference and amoralism (Zangwill 2008).

The internalist theory cannot simply rule out certain cases or kinds of agents on an ad hoc basis to diffuse the challenges to the theory. They must have a well-defended reason for setting the particular conditions that they do (Svavarsdóttir 1999 p. 179)⁵⁵

⁵⁵ As Bjorklund et al., point out, whatever the exclusion or application conditions run the risk of making internalism "explanatorily impotent" or "vacuous" (Bjorklund et al. 2012 p. 128). For those arguments see: (Lenman 1999 p. 298-9; Roskies 2003 p. 53-55; Miller 2003 p. 221).

2.2 Amoralists do not make moral judgments

The most common internalist response to amoralism is to argue that amoralists do not make judgments at all.⁵⁶ The locus classicus of this response is found in R.M. Hare's discussion of moral cynics. According to this internalist response, the amoralist is not making the moral judgment at all; he is simply stating the fact that *other* people judge, or it is conventionally accepted, that the act in question is morally wrong. In Hare's language, that is, amoralists use moral language in merely an inverted comma sense (Hare 1952 p. 125). Those moral terms have a different meaning when they are used in this way.

The Hare-style strategy is a common non-cognitivist one. Since non-cognitivists claim that moral statements are expressions of non-cognitive, affective states, they are always motivational. So if the agent is not motivated to act on his purported moral judgments, this means they cannot be making a sincere moral judgment at all. Hare's original comments provide the general form of the argument that internalists use to defend their view against the amoralist counter-example.⁵⁷

In order to maintain the amoralist as a counter-example to internalism, externalists have to give a counter-response to Hare's inverted comma analysis.

⁵⁶ A related kind of response to the amoralism objection to internalism is to re-describe what the amoralist is doing as a rejection of the mainstream system of morality altogether in favor of an alternative value system: "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 no value" (McNaughton 1988 p. 140; Sayre-McCord 1997 p. 64-65).

⁵⁷ Jonathan Dancy similarly argues that internalists should question the way externalists have described the phenomenon of amoralists. The amoralist is "person who sees the institution of morality from the outside, as something whose claims on us he rejects. But this person does not accept the moral judgments whose relevance he denies; at best, he merely knows what judgment would be made by others" (Dancy 1993 p. 5). A related point is made by Brook Sadler, who points out that perhaps "some amoralists are amoral in virtue of the fact that they promote new, non-moral evaluative standards" (Sadler 2000 p. 119). She uses this point to show that such amoralists "are untouched by the inverted-commas argument" (Sadler 2000 p. 119).

Externalists can either respond to the argument at a general level, or to specific versions of the argument. I am going to give three general responses first. I will then respond to the specific ways that Smith and Jesse Prinz make use of a Hare-style argument in defending their versions of internalism.

There are three general arguments against the Hare-style response to the amoralist. First, I present an argument based on moral disagreement (Brink 1989; Copp 1995; Sturgeon 1986 p. 120-124). Second, I raise the response that the Hare-style response does not take the amoralist challenge seriously enough (Brink 1989). And third, I question whether or not the inverted-commas analysis of amoralists' moral judgments can in fact ward off the externalist threat.

If we assume that the inverted comma analysis of the amoralist is correct, this yields some puzzling conclusions about moral disagreements between amoralists and non-amoralists. This argument unfolds in two parts. While the normal agent expresses the belief that 'X is morally wrong,' according to the inverted-comma analysis, the amoralist expresses instead the belief that 'other people think that X is morally wrong.' This means they express two different beliefs (Copp 1995 p. 13). When someone denies what the amoralist says, he merely denies that *other* people think that one ought to give to famine relief, for example. This means that when the amoralist and the normal person seem to have a moral disagreement about whether or not something is morally obligatory, this disagreement is merely an illusion (Copp 1995 p. 13). In some sense, a non-amoralist *cannot disagree* with an amoralist. This is a puzzling result since

it seems that we can engage in genuine moral debate with people like the amoralist, who lack motivation to act on their moral judgments.⁵⁸

The second, part of this puzzle has to do with the sense in which we can actually be disagreeing with the amoralist on a non-cognitivist view. The non-amoralist and the amoralist Roger, for example, may both say that 'killing babies for fun is wrong' (Brink 1989 p. 85). Further, both may explain why killing babies for fun is wrong, and share reasons for asserting this, such as the pain that it causes the baby and his or her family, or the disrespect for human rights that the action exhibits. The amoralist's statement, however, is not accompanied by any corresponding negative attitude towards or motivation to abstain from killing babies for fun (Brink 1989 p. 86). According to the Hare-type analysis of the amoralist, the amoralist is only parroting what is conventionally accepted as wrong.

Based on the non-cognitivist's semantic analysis however, "[i]t would seem to follow ...that the amoralist and the moralist must be in *moral disagreement*, because they hold different attitudes towards the same actions, for example" (Brink 1989 p. 86). So non-cognitivist should be able to say that the amoralist and the non-amoralist do disagree because they hold different *attitudes* toward killing babies for fun. But this does not accurately represent the difference between the amoralist and the non-amoralist either, the amoralist and the moralist are not disagreeing about first order ethical principles, or ethical theories, or even about "what makes good or right actions

⁵⁸ Copp goes on to argue that Hare could be right about the following claim about the meaning of moral terms: "Hare's claim is that the fact that a person using moral terms to make a moral judgment expresses a corresponding conative state is due to the *meaning* of moral terms when they are used literally. He is also claiming that this aspect of the meaning of moral terms can be canceled or eliminated by placing the terms in "inverted commas." Copp argues that this view could still be compatible with moral realism and externalism, according to Copp's unique "realist expressivism," (Copp 1995).

good or right” (Brink 1989 p. 86). The content that they are disagreeing about remains mysterious.

The second general response to Hare-style analysis of the amoralist is the charge that it dismisses the amoralist too easily, or in Brink’s words it “doesn’t take the amoralist’s challenge seriously enough” (Brink 1989 p. 47). Brink sees the amoralist as presenting a classic kind of moral skepticism, sometimes called practical moral skepticism. It would be quite surprising if this philosophical challenge could simply be dissolved by claiming that such a skeptic is not really using moral language the way the rest of us normally do.⁵⁹ Taking the amoralist challenge seriously, then involves entertaining the possibility that some people are not motivated by their moral judgments, and that they may ask why they should be.⁶⁰

Brink explains that all non-cognitivists must ignore what the amoralist says in a certain way, “by accusing the amoralist of a conceptual or semantic confusion” (Brink 1989 p. 84). The Hare-type analysis of the amoralist gives us only a “skeptical solution” to a skeptical problem. The skeptical problem is the existence of someone who is not motivated by his or her own moral judgments. The existence of this person does not worry the non-cognitivist, because for them, “the very meaning of moral judgments

⁵⁹ Brink says two distinct things about the skeptical challenge of the amoralist: “if we are to take the amoralist challenge seriously, we must attempt to explain why the amoralist should care about morality” (Brink 1989 p. 48). And “amoralist skepticism” is “skepticism about the justification or rationality of moral demands” (Brink 1989 p. 48). The question of why someone should care about morality is ambiguous, and brings us close to the question of the authority or normativity of morality. Brink’s response to internalists can be strengthened if these two issues were disambiguated. Why and how we are motivated to act on our moral judgments is distinct from the question of whether or not we should care about moral demands or give priority to our moral judgments in our decision making.

⁶⁰ When they ask why they *should* be motivated to act on their moral judgments however, they are asking a normative question. One potential answer to that question has to do with rationality, but there are certainly other possibilities. But explaining to the amoralist that it is irrational not to care about morality is a possibility open to a cognitivist externalist that is not open to a non-cognitivist or most internalists according to Brink (Brink 1989 p. 84-5).

ensures motivation, the skeptical problem posed by the amoralist is, according to the non-cognitivist, really misconceived” (Brink 1989 p. 84).

The third response, examines what the inverted-commas analysis of amoralist moral judgments can show. Since Smith and other internalists rely heavily on Hare’s inverted-commas analysis as a way of explaining the way that amoralists make moral judgments, it is important to start by looking at what Hare means by this phrase and whether or not it supports internalists’ argument that amoralists do not really make moral judgments at all (Sadler 2000 p. 114).

We noticed that it is possible for people who have acquired very stable standards of values to come to treat value-judgments more and more as purely descriptive, and to let their evaluative force get weaker. The limit of this process is reached when, as we described it, the value-judgment 'gets into inverted commas', and the standard becomes completely 'ossified'. Thus it is possible to say 'You ought to go and call on the So-and-sos' meaning by it no value-judgment at all, but simply the descriptive judgment that such an action is required in order to conform to a standard which people in general, or a certain kind of people not specified but well understood, accept (Hare 1952 Sect. 11.1)

Sadler argues that in order to understand how useful inverted-commas analysis can be for the internalist, we have to have a little background on Hare’s purpose for using this phrase.

There are three things we know from Hare about the inverted commas sense of a word: (1) that it reflects or effects a change in standards; (2) that it represents a purely descriptive use of a word that also has, or primarily has had, an evaluative sense; and (3) that there is a connection between the inverted commas sense of a word and convention (Sadler 2000 p. 115).

What is important to note here is the contrast between the purely *descriptive* use of the word in the inverted-commas sense and the *conventional* use of the word. The conventional use of the word must “have an evaluative sense in order for it to be apparent that the evaluative sense is not being invoked when the word is used in an

inverted commas sense or purely descriptive sense” (Sadler 2000 p. 115). This is important for the internalists analysis of the amoralist because it means that the amoralist making an apparent moral judgment, or using moral language in an inverted commas sense, has to be able to “*allude to* the conventional, evaluative sense” of the moral terms being used (Sadler 2000 p. 116).

In order for the amoralist to be able to “allude” to the evaluative sense of the word or judgment, he or she must have certain capacities. Sadler argues that these capacities include the ability to have “a real understanding of the conventional evaluative sense” of the moral terms that they are using in their moral judgments (Sadler 2000 p. 116). An easy way to misunderstand this argument is to take it to be saying that in order for the amoralist (or anyone) to *utter* a sentence that contains a term that has an evaluative sense, they must have a robust understanding of the evaluative sense of the term. This cannot be correct. Surely we utter terms and even *allude to* things that we do not understand all of the time. This is not the claim being made. Instead, it is that in order for the amoralist to make a moral judgment in the inverted-commas sense, he or she must be able to understand the evaluative sense of the moral terms. In other words, if someone used moral terms in an utterance and truly did not understand that those terms have an evaluative sense, this would not even count as using them in an inverted-commas sense (at least in the way that Hare intends the significance of this term). Such a use would be some other kind of use, if it could even be called a use at all, of moral language.

If this analysis is correct, then it turns out that after all, saying that the amoralist’s moral judgments are just inverted commas moral judgments does not tell us very much

about those judgments or why we should think of them as “not real moral judgments” (Sadler 2000, p 116). This is because “it is hard to see just what is missing from his judgments which serves to classify his [amoralist’s] judgments as not *real* moral judgments” (Sadler 2000 p. 116).

It seems that part of what the internalist is hoping to get at when they say that the amoralist does not make real moral judgments, is that the amoralist “alludes to a value judgment [but] does not herself *endorse* it” (Sadler 2000 p. 116). But as Sadler points out, if this is the problem with the amoralist’s inverted commas use of moral judgments, then this does not seem to be sufficient to illustrate that the amoralist is exhibiting “some kind of confusion or incoherence” as Hare intends (Sadler 2000 p. 116). “We can, though we are not amoralists, make judgments about what other people judge to be right and wrong, without endorsing those judgments” (Sadler 2000 p. 116). In fact we do this all of the time and it does not lead us to think we are confused or incoherent.

How then should we understand what Hare means when he “says that when we make inverted commas judgments we are not ‘making a value judgment ourselves, but alluding to the value judgments of other people’” (Sadler 2000 p. 117)? Sadler argues that alluding to other people’s moral judgments “requires the agent to *make moral judgments*.” In fact, someone who is able to allude to other people’s moral judgments, in Hare’s sense, has to be able to make two distinct types of moral judgments, both first and second order judgments (Sadler 2000 p. 117-118). If the agent is an amoralist, the externalist contends, they are able to make moral judgments but fail to be motivated by them. On Sadler’s analysis, this phenomenon involves two distinct moral judgments.

[A]moralists...first...exercise their ability to make first-order judgments which concur with the first-order judgments of normal moral agents; and second, they

make some kind of second-order judgment by which they reject the evaluative sense of the first-order judgment” (Sadler 2000 p. 118).

We can conclude from this reevaluation of Hare’s inverted commas sense of moral judgments that simply citing inverted commas as an explanation of the moral judgments of amoralists is not a useful internalist rebuttal to externalism. Internalists who rely on inverted commas must be more precise about what element of the amoralist’s moral judgment is missing that disqualifies their judgment from being like our moral judgments (Sadler 2000 p. 117). Of course, what is missing cannot simply be motivation. If that is all the internalist has to offer then they are simply assuming the truth of internalism.

2.2.1 Smith’s Response to Amoralism

To be able to maintain that there is a necessary, inherent, or essential connection between making a moral judgment and being motivated by it, internalists have tried to respond to amoralist counter-examples with Hare-style interpretations of their moral judgments. Smith also gives a Hare-type response to the externalist challenge of the amoralist; the amoralist “tr[ies] to make a moral judgment but fail[s]” (Smith 1995 p. 68-9).

What is unique about Smith’s argument against the possibility of amoralism is an analogy with color judgments. Imagine someone who is born blind but has been attached to a machine her whole life that allows her to reliably make color judgments based on the machine’s translation of the presence of surface reflectance properties into tactile sensations. According to Smith, we would not say of this blind person that she is making color judgments in the same sense that sighted people do, even though she can reliably detect the presence of colors. In the same way, we cannot say that the

amoralist is making moral judgments, even though he may be able to reliably detect moral properties (Smith 1995 p. 69).

Smith's argument based on the analogy with color judgments fails on three grounds. First Smith does not give us a reason to think that color judgments are an appropriate analogy with moral judgments, so that color judgments can reveal the nature of moral judgments (Brink 1997 p. 25). Every argument from analogy has to start with a useful analogy. There are a considerable number of differences between the color judgments and concepts and moral judgments and concepts. For example, part of the analogy involves comparing our visual perceptual system with our motivational system. Smith certainly does not mean to suggest that we detect moral properties in the world in an analogous way to the way in which we perceive colors. But if he is not adopting that type of view, it is not at all clear why we should think that these two different systems function in ways that are similar enough to be able to draw informative conclusions about one from facts about the other. This is enough to create doubt about how much this argument could show, even if it was successful.

At the same time, Smith's characterization of the nature of color concepts and their relationship to visual experiences is by no means uncontroversial in the philosophy of perception or mind. The obvious question to ask Smith here would be: what is a color concept in the first place? Whether or not color concepts even exist and in what sense is contested. See (Tye 2002; Rosenthal 2001, 2010; Papineau 2006; Mandik 2012). Whether or not color experience is needed in order to have knowledge of colors is also an area of ongoing debate. See (Unger 1966; Lewis 1990; Alter 1998; Stoljar 2005; Jackson 1986).

Second, Smith does not provide a good reason to accept the claim that “certain visual experiences are constitutive of color concepts” (Brink 1997 p. 24). It is perfectly reasonable to think that though a blind person might lack “visual experiences,” they could still use color terms in a meaningful way and issue judgments about the colors of objects, for example, based on what they have learned in other ways (Brink 1997 p. 24). Railton’s character of the amoralist Roger, seems like an obvious example of a person who uses moral concepts, and makes moral judgments, reliably detecting moral properties, without any corresponding motivation, just like the blind person issuing color judgments, using color concepts (Railton 2009 p. 166).

Third, Smith’s view implies that those who become amoralists or morally indifferent mid-life, lose their ability to use moral concepts and make moral judgments, and can now do so only in the inverted comma sense. If visual perception is necessary for using color concepts and making color judgments then we would have to say that those who become blind in mid-life suddenly lose their color concepts and ability to make color judgments. We have little reason to think that people who become blind later in life, those who are temporarily blindfolded, those who simply have their eyes closed, or those who have only night-blindness or color blindness lack color concepts and cannot make color judgments consider (Brink 1997 p. 24; Roskies 2003 p. 60).⁶¹

This problem for Smith’s response deepens when we consider the empirical evidence from acquired psychopathy. Roskies and Damasio have found that patients with acquired psychopathy, as the result of damage to the Ventromedial frontal region

⁶¹ ...the principled amoralist need not be congenitally amoral...They made moral judgments with motivational effect prior to the onset of depression or neurological damage. They continue to make the same discriminations and judgments using moral terms but now without motivational effect” (Brink 1997 p. 24).

(VM) of their brains can be characterized as to some extent having lost their moral motivation and their ability to act on their moral judgments. However, they have limited other cognitive problems and seem to continue to make the same moral judgments that they made before the injury (Roskies 2003, 2006; Damasio et al. 1994; Roskies 2008). Roskies describes them as “lacking motivation” (Roskies 2003 p. 63). A patient who sustains such a brain trauma later in life and subsequently loses his or her ability to be motivated by his or her own moral judgments can be thought of in a similar way to the person who becomes blind later in life (Roskies 2003).

2.2.2 Prinz’s Amoralist

Jesse Prinz uses empirical psychological research to offer a new version of the classic Hare-style response to the amoralist counter-example to internalism. Prinz uses psychopaths as his example of real life amoralists. Psychopaths are more prone to violence, immoral behavior, and criminal behavior than non-psychopathic individuals (Glenn et al. 2009 p. 497-505; Belmore and Quinsey 1994; Ishikawa et al. 2001). Yet, they seem to be able to make the same moral judgments about the wrongness of these actions as non-psychopaths do (Cima, Tonnaer, and Hauser 2010 p. 59-67).

Prinz’s theory of epistemic emotionism involves a commitment to internalism and so Prinz gives an explanation of the moral psychology of psychopaths that is compatible with internalism (Prinz 2007 p. 46). Prinz argues that psychopaths do not actually possess the same moral concepts that the rest of us do. Instead, Prinz likens psychopaths to anthropologists who are merely reporting on what other people take to be morally right and wrong; Prinz writes, “the concepts that psychopaths express when

they use the words 'right' and 'wrong' differ from our concepts in both sense and reference. They can mention these concepts, but they can't use them" (Prinz 2006 p. 38).

To support his interpretation of the moral judgments of psychopaths, Prinz relies on two pieces of evidence. First, psychopaths are less sensitive than non-psychopaths to the difference between moral rules and conventional rules (like etiquette)⁶² (Blair 1995; Prinz 2007 p. 44). This work shows that inmates who scored high on psychopathy checklist rated violations of conventional rules to be very serious and impermissible even if society and authorities said that it was all right to break the rules.⁶³ The second piece of evidence is that psychopaths suffer deficiencies in a range of emotions, for example, they exhibit a lack of fear and sadness and a lack of a "vicarious stress response" when they see other people in pain (Prinz 2007 p. 45). Prinz argues that normal moral concepts are "essentially emotionally laden" (Prinz 2007 p. 46). Because psychopaths do not possess the same moral concepts that we do, they are not able to issue genuine moral judgments. Instead they merely "report on morality" (Prinz 2006 p. 38).

Prinz's empirical version of the Hare-style response to the amoralist counter-example is vulnerable to two criticisms. First, Prinz does not offer a justification for thinking that being able to make the moral conventional distinction is essential to possessing a moral concept or the ability to make a moral judgment (Roskies 2008 p. 202). An independent argument, which Prinz does not provide, is necessary for thinking

⁶² Similar to an argument made against Roskies (Kennett and Fine 2008).

⁶³ Patients with damage to their prefrontal cortex late in life do, however, seem to be sensitive to this distinction (Saver & Damasio 1991).

that the moral conventional distinction is an essential part of having moral understanding or possessing moral concepts (Roskies 2008 p. 202).

Second, the externalist does not have to deny that competency with moral concepts, and possession of a standard range of moral emotions, is *linked* with motivation. Instead, they deny that there is a necessary connection between the two, or that moral judgments are necessarily or intrinsically motivating. The co-occurrence of emotional and moral deficits cannot establish that the emotions are necessary for possession of the moral concepts.

Prinz's Hare-style interpretation of the amoralist work less well when applied to cases of occasional amoralism or diversity in moral motivation. Agents who are occasionally amoral do seem to possess normal moral concepts, but still fail to be motivated by their moral judgments (Bagaric 2002). These are agents that are otherwise mentally healthy, in general are motivated by their moral judgments, and often act consistently with them. Yet they may sometimes fail to be motivated by them or at least are motivated by them to different extents at different times (Zangwill 2008 p. 101). In these cases, Prinz cannot claim, as he did with psychopaths, that the agents fail to possess moral concepts necessary to make moral judgments, possession of a concept does not seem to come and go that easily.

Prinz could respond to the occasional amoralist case by agreeing that the moral judgments agents make in cases of occasional amoralism do involve standard moral concepts (Bagaric 2002). Although the agents lack motivation when they make some moral judgments, they must have a "disposition" to be "motivate[d] under the right circumstances" (Prinz 2006 p. 38). On one reading, this response means that Prinz

does not hold anything like a strong form of internalism. As he does not maintain that moral judgments necessarily or inherently motivate. If this is not the correct interpretation, it seems that the challenge of the occasional amoralist puts the burden of proof on Prinz and other internalists to explain why motivation is activated in some cases but not in others. Externalists offer a straightforward explanation of the phenomenon of occasional amoralism. Moral judgments may be motivationally inert until the agent also possesses the appropriately related desire or other conative state (Svavarsdóttir 1999 p. 170).

2.3 The Impossibility of Community Amoralism

In response to the counter-example of the amoralist, internalists attempt to save their theory by reinterpreting the amoralist's moral judgments along the lines of Hare, claiming that they aren't really making moral judgments at all. As we have seen, this interpretation is not convincing. The response to amoralism discussed in the following section concedes that individual amoralists might seem like counter-examples to internalism. However, we can conceive of individual amoralists only because our overall moral practice is consistent with internalism. In other words, amoralists are able to make what appear to be genuine moral judgments because they are parasitic on our community moral practice. We cannot even conceive of a community of amoralists. And this, the response continues, shows us that the inverted comma interpretation of the moral judgments of the amoralist is vindicated (Blackburn 1998 p. 61).

James Lenman argues that while we can conceive of amoralists who make genuine moral judgments, we cannot conceive of "global amoralism" (Lenman 1999 p.

144). Lenman writes, “you can’t be an amoralist *on your own*...You can’t be a *group of amoralists on your own*” (Lenman 1999 p. 445). Lenman imagines what he calls a “preposterous story” about a planet populated entirely by amoralists (Planet Amoralism), who are very good at detecting and studying moral properties, and reporting on the moral facts, but unlike us, “they don’t care about them...do not allow their discoveries about what these facts are to impinge in any way on their habits or desire and action” (Lenman 1999 p. 446). Lenman argues that if there was only a contingent connection between making moral judgments and being motivated by them, then this story should not strike us as strange at all. If internalism is true it will. Since it does strike us as absurd, then this is evidence in favor of internalism (Lenman 1999 p. 446).

Lenman’s argument is supposed to support the internalist rejection of the possibility of amoralism. However, a crucial part of his argument is that the individual amoralist makes moral judgments only in an inverted comma sense. Lenman goes on to say that this is what makes the planet of amoralists preposterous. On planet Amoralism: “... there is *nothing* for the content of the judgment that something is ‘morally good’ (is what other people think is morally good) to be” (Lenman 1999 p. 452).

So on planet Amoralism the moral judgments of its denizens lack propositional content (Lenman 1999 p. 452-3; Bedke 2009 p. 193). This is because, according to Lenman, the best any amoralist can be doing is reporting on the genuine moral judgments of other people. But on the planet of amoralists, there are not any genuine moral judgments being made to be reported on. Of course, the externalist disputes this interpretation. Because Lenman assumes the inverted comma interpretation of the

amoralists' moral judgments from the very beginning it is hard to see what his thought experiment can actually add to the internalist's defense of their view.

A related kind of community amoralism is found in M.S. Bedke's thought experiment about the planet Amoralville (Bedke 2009 p. 194-5). On Amoralville people perform actions that seem identical to the actions that we call moral actions here on earth. But on Amoralville no one uses moral language or concepts, nor do people make moral arguments. Instead, people do what they do based on judgments of their own prudential interests. On this planet, a strong ruler doles out swift and severe punishments for breaking a set of rules (that happen to coincide with our moral rules). Bedke stipulates that no one on Amoralville is motivated by anything but self interested fear of punishment (Bedke 2009 p. 194). Due to radio transmissions from earth making it to Amoralville, the inhabitants are able to hear us using moral language and making moral arguments:

With the introduced moral vocabulary, the residents of Amoralville learn to apply moral concepts correctly. As a result, Amoralville residents correctly pick out what is right and wrong, acknowledge obligations, and can correctly categorize that which they (morally) ought to do. In fact, forming first personal putative moral judgments and speaking in ethical terms becomes kind of a fad in Amoralville, though the judgments never garner any motivational force, and moral demands simply do not weigh with them. Residents of Amoralville are at all times solely motivated by their own interests (Bedke 2009 p. 194-5).

Given this characterization and the assumption that the inhabitants of Amoralville are able to acquire linguistic competence with moral terms, Bedke remains convinced that we would still not characterize these people as making moral judgments. This is because, Bedke argues, on Amoralville morality has lost the "action-guiding character" that it possesses here on earth (Bedke 2009 p. 195).

While this imaginary case of community amorality maybe better articulated than in Lenman's version of community wide amorality, it still fails on two grounds. The first problem is that it draws the unsupported conclusion that the citizens of Amoralville fail to make moral judgments. It remains open to the externalist to say that the citizens of Amoralville do make moral judgments; they just lack the related emotions, desires, or other conative states to act on them.

Second, the externalist does not have to reject the claim that morality is action guiding when she rejects the internalist notion that there is a necessary, inherent, or essential connection between moral judgments and motivation. For externalism, moral discourse continues to be action guiding in the sense that moral actions often arise as a result of moral discourse and that moral beliefs, when coupled with other propositional attitudes, *often* produce moral actions. This is not the same as claiming, as Bedke seems to, that being action guiding is an essential feature of making a genuine moral judgment.

The idea that morality is fundamentally action guiding may alternatively have to do with the authority of morality. Part of my response to this case is beyond the scope of this chapter and will be addressed in Chapter 4 on the authority objection to realism. In short, morality itself can be action guiding, without action guiding-ness being an essential feature of making a moral judgment. We are morally obligated to act on our moral judgments. We may also be morally obligated to be motivated to act on our moral judgments or to attempt to develop the disposition to be motivated to act on our moral judgments. But this does not mean that when we make a moral judgment, for it to count as such, it has to be acted on, or dispose us to act on it, by itself.

2.4 The Amoralist Is Motivated

Internalists may deny that the morally indifferent actually fail to be motivated by their moral judgments. Internalists claim that externalists are describing a psychological profile that cannot exist. If the hypothetical agent insists that they are not motivated by their moral judgments, an internalist might insist that the agent is simply confused with respect to his or her own moral commitments or motivations (Björnsson 2002 p. 336; Shafer-Landau 2000 p. 152).⁶⁴

While there is no doubt that we have imperfect access to our own motivational states and that introspection has its limitations, it is uncomfortable for the internalists to have to explain a wide variety of cases in this way. Although amoralism is the most commonly discussed form of moral indifference, this response must also stand when applied to otherwise normal agents who engage in occasional amoralism and simple diversity in moral motivation. On this response all of these people are mistaken about their level of motivation or its absence. The other curious thing about this kind of response is that it only comes up when the agent introspects and finds they lack motivation, not when internalists introspect and find that motivation always accompanies their moral judgments!

⁶⁴ Björnsson has argued that the phenomenon of “silencing” can explain the way in which it might sometimes seem (even to the agent) that he has no motivation to act on his moral judgments (Björnsson 2002). Using the example of the judgment that he has an obligation to report some income to the IRS, X explain how silencing differs from the externalist explanation of these phenomenon, that is, that the agent had no motivation to act on his moral judgment. The difference is that silencing is consistent with internalism. Silencing means that the agent does not feel and internal struggle, does not experience guilt, or anxiety, and is not disposed to act in ways that someone in a moral conflict would act. Instead the agent is not aware that they have a counter veiling motivation to act on their moral judgment, because it has been silenced by their opposing optation to preserve their self interest. X argue that “the moral optation is silenced by the personal cost” of acting on the moral judgment. (Björnsson 2002 p. 338). The same thing happens in cases when the agent feels like in a world where everyone else is behaving immorally, their own moral actions can make little difference, and thus seem to lose all motivation to act on their moral judgments (Björnsson 2002).

It is certainly plausible that agents can make these kinds of mistakes, the internalist cannot simply stipulate that the agents are mistaken as a way of saving the theory from these counter-examples without “any detailed acquaintance with the psychology of the agent” (Shafer-Landau 2000 p. 152). Shafer-Landau likens this strategy to the way a defender of psychological egoism may re-explain any instance of altruistic motivation as actually self-interested (Shafer-Landau 2000 p. 152; Zangwill 2008 p. 104). Simply asserting that a hypothetical agent must be morally motivated cannot get the internalist very far.

Bjornsen and Olinder argue that internalists commit the “Deviance Fallacy” and that this explains why internalists think it is impossible that an amoralist, a person who makes moral judgments but is not motivated by them, could exist. Internalists “have unconsciously relied on the commonsense assumption that actual moral opinions normally are followed by genuine moral motivation” (Björnsson and Olinder 2012 p. 10). In other words, internalists only consider the amoralist against a background of normal agents who are routinely motivated to act on their moral judgments. With this background in mind, amoralists “will seem significantly deviant” (Björnsson and Olinder 2013 p. 10). The deviance fallacy involves:

[F]ailing to recognize that the intuitions about such scenarios depended on the unconscious background assumption, we committed the ‘deviance fallacy’. Mistaking an intuitive judgment of deviance from our model of paradigmatic cases for a judgment of conceptual impossibility (Björnsson and Olinder 2013 p. 11).

They use the “cynical hypothesis” to argue for the existence of the deviance fallacy. The cynical hypothesis says that “the tendency for people to act in accordance with their moral opinions ultimately stems from a desire to appear moral”(Björnsson and Olinder

2013 p. 1). The point of the cynical hypothesis is to “bracket our ordinary understanding of moral psychology.” By doing so, we have the opportunity to reevaluate the intuitions that lead to an acceptance of internalism, specifically strains of internalism that involve the claim that amoralism is a “conceptual impossibility.”

3. Pluralistic Externalism and Moral Indifference

The pluralistic externalism that I argued for in the previous chapter, does not assume that there is one dominant or uniform type of interaction between moral judgments and moral motivation. The fact that we are often motivated to act on our moral judgments is a purely contingent matter. Pluralistic externalism provides sufficient conditions for being motivated to act on one’s moral judgment but it does not provide necessary conditions for being motivated to act on one’s moral judgment. For example, on the pluralistic view, sometimes an agent may be motivated by the desire to do the right thing. Other times, it may be that the agent’s sympathy for another person causes them to act on their moral judgment.

At the same time it does not make any empirical claims about what the breakdown of actual human motivation comes from, in any one individual or across the human population this may vary. Because on this view agents are motivated to act on their moral judgments by a wide variety of sources, there are also a great many ways this motivational system may fail to function. In this chapter I have discussed several forms of moral indifference, various kinds of amoralism, and listlessness.

These forms of variability in the relationship between moral judgment and motivation serve as counter-examples to internalism, which insists there is a necessary,

inherent, or essential connection between making a moral judgment and being motivated by that judgment. Internalists defend their position centrally by reinterpreting these forms of moral indifference. Their most popular strategy is to claim that the morally indifferent fail to make genuine moral judgments in one way or another. What all of the internalist responses have in common is that they rely on internalist assumptions about what it means to have a moral belief, make a moral judgment, or possess a moral concept. In other words, people who are defending the necessary connection between moral judgments and motivation, define moral judgment to fit their theory, then of course when they look counterexamples to the theory, they will say that there aren't any moral judgments being made in the counter-examples.

3.1 An Objection to Modal Intuitions and a Response

Some have argued that while the range of situations and characters that I describe in this chapter might be entertaining, we should be very skeptical about how much they can actually illuminate the relationship between moral judgments and motivation (Björnsson 2002 p. 331-2). In other words, I can imagine all sorts of strange and puzzling states of affairs, and psychologies, so what? (Björnsson 2002 p. 331-2). I am sympathetic to this objection and also think we should be skeptical about the extent to which twin earth thought experiments, for example, can tell us about metaphysics. Fortunately, these aren't imaginary cases in the relevant sense. The characters and phenomena that this chapter presents are almost all just folk psychological descriptions of our moral lives, those of the philosophers cited, and those of people we know and observe.

These critics will also point out that there is considerable disagreement in intuitions about how to interpret the hypothetical case of amoralism and instead we should focus on finding the theory with the greater explanatory power (Björnsson 2002 p. 332) This is another point I concede, bare clashes of intuitions about imaginary cases cannot settle the dispute between internalism and externalism. We have to consider which theory taken as a whole offers better explanation of the range of phenomenon in question. Of course, when considering competing theories in moral psychology, we must also consider the empirical evidence on both sides. Evidence from acquired psychopathy and VM patients is not definitive at this point for either internalism or externalism. I have argued in this chapter that pluralistic externalism is the theory that provides the better explanation of the phenomenon of moral indifference.

Chapter 4: The Authority and Normativity Objections

In Chapter 1 I presented family of classic queerness objections that have been posed against moral realism, and focused on one strain of that argument called the normative force objection. I argued that the normative force objection is ambiguous between two interpretations, the internalism objection, and the authority objection. In Chapters 2 and 3 I responded to the internalism objection. In Chapter 2 I offered arguments that undermine considerations that are often taken to support internalism and defended the alternative view, externalism, against criticisms. I also presented a version of externalism, pluralistic externalism, which is best suited to explain the relationship between moral judgment and motivation. Chapter 3 extended the response to the internalism objection by focusing on the phenomenon of moral indifference and amorality. In it, I argued that these phenomena are also more easily explained by a pluralistic externalist theory of moral motivation than by internalism or competing versions of externalism.

1. The Authority Objection

This chapter focuses on responding to the authority objection. According to the authority objection, moral realism, (especially *naturalist* moral realism) wants to have its cake and eat it too. Morality, according to realism, consists of an objective realm of facts. On this view, obligations and duties exist, regardless of what beliefs people have about them or emotive reactions they have to them. Naturalist moral realism, in particular, insists that these facts and properties do not introduce anything mysterious

that we should worry about into our ontology, nor do they commit us to anything epistemologically troubling, like a special faculty of intuition with which to detect them. Instead, moral properties and facts are just like historical, sociological, biological, or psychological facts. We have access to them in the ordinary way, through observation. The kind of relationship they have with more basic physical facts and properties can be understood as one of identity, constitution, or supervenience.

But, the objector to realism protests, morality is not like history or psychology. The facts of history and psychology aren't facts that have any authority over us. They do not have to matter to us in the way morality does. They are not action guiding or practical. I can shrug and walk away from them in a way I cannot shrug and walk away from moral facts and properties. The authority objection questions the authority that any set of facts or properties could have over our behavior. In other words, the moral realist seems to lack a convincing response to the moral skeptic who asks, 'why should I be moral?' The surprising upshot of this objection is that, while the realist may be able to secure the objectivity of moral rules and the truth of moral facts, they cannot secure its authority.

This objection is cashed out differently depending on whether the objector is focusing on moral realism as a whole, as Christine Korsgaard does in her classic normativity objection to moral realism or if the objector is focusing specifically on naturalistic versions of moral realism, as Derek Parfit does in his recent normativity objection and triviality objection to what he calls non-analytical naturalist moral realism.

The authority objection charges that [naturalist] moral realism cannot explain the authority morality has over us.⁶⁵ The idea of the authority of morality is vague. It is broader than normativity, but includes normativity. It may specifically mean the inescapability of moral requirements, the special level of importance that morality has in directing our decision making and life choices, the way in which moral considerations seem to override all others, or the categorical nature of moral reasons.⁶⁶ Without narrowing down authority at all, the following is the basic argument from authority.⁶⁷

- (1) Morality has a certain kind of authority over us.
 - (2) Naturalist moral realism precludes this kind of authority.
- Therefore, naturalist moral realism must be false.

Two obvious strategies emerge for responding to this argument. First, the realist can deny premise one, arguing that, in fact morality does not have the kind of authority that critics claim it has. Second, the realist can give an account of the authority that comports with naturalist moral realism.

My strategy will employ both of these responses. Premise one is true, given a certain, limited account of authority. The first step here is to consider what a critic might mean by the authority of morality. I consider two plausible accounts of morality's authority. The first type of authority I consider is authority as normativity, according to Korsgaard's rationalist neo-Kantianism. I present her normativity based objection to moral realism and respond to it. Second, I consider Parfit's recent account of normativity.

⁶⁵ One of the earliest sources on the authority of morality is Joseph Butler (Butler 1726 Sermon 2-3).

⁶⁶ Sometimes the authority of morality is also explained in terms of the notion that "moral requirements provide reasons for acting which override other kinds of reasons" (McNaughton 1988, p. 115).

⁶⁷ Other ways you could put this in more detail—on naturalist moral realism moral properties or facts are similar to the kinds of properties or facts that don't have authority over us, like facts about history.

Contrary to their accounts of authority, I argue for a minimal authority for morality rooted in objectivity. Given the minimal account of authority, premise two is false, this minimal kind of authority is no more a problem for naturalist moral realism as a metaethical theory than any other metaethical theory.

1.1 What has Authority?

Before discussing what authority might plausibly mean for naturalist moral realism and what it means in the context of Korsgaard's and Parfit's influential objections to moral realism and naturalistic moral realism, respectively, it is important to make two crucial distinctions in what we are talking about when we talk about the authority of morality.

First, it is important to distinguish between the authority of morality itself, that is of moral rules, moral facts, or moral properties, on the one hand and the authority of moral language and terms on the other hand. Some claim that the terms "good" or "ought" have a kind of force that comes from the role they play in our language. Copp refers to this type of authority as "generic normativity" and explains that this type of normativity has to do with the "semantic connection" that moral claims gave "to decisions about action or choice" (Copp 2007). Copp takes this form of normativity to be difficult for the realist to explain because of the way realists understand moral language or moral claims as "express[ing] propositions that attribute moral properties;" they are descriptive claims (Copp 2007). The puzzle for a realist is how a claim that describes a part of the world can also "evaluate, recommend, or prescribe" at the same time (Copp 2007).⁶⁸

⁶⁸ Copp's explanation of generic normativity comes from his "standard-based account" (Copp 1995 p. 267).

This chapter will focus on the authority of morality itself, rather than the authority of moral language, specifically whether moral facts or properties carry with them an authority over us that other kinds of properties or facts lack, rather than whether moral language has a prescriptive or normative force.

Second, it is important to distinguish between the idea that morality itself, the properties of goodness, rightness, wrongness, or the moral facts or truths *themselves* have authority over us and the claim that our *beliefs* that have moral content have authority over us. This is a distinction that is often overlooked. If it is the moral facts themselves that are supposed to have authority over us, this is something different from saying that psychologically speaking our moral beliefs, happen to (or often do) have a certain kind of priority in our decision making. I think that this is one of the sources of the ubiquitous conflation of the authority objection with the internalism objection. Another source of the conflation might be the distinction between reasons for action and motivational reasons. If it is the moral facts or considerations themselves that are supposed to have authority, then we are saying that they give us reasons for action (it does not have to be action necessarily that is just the easier way to put it). If we are saying that it is our moral beliefs or judgments that have the authority then we may be talking about reasons that are going to motivate us in some way (See Chapter 2).

1.2 What is Authority: Inescapability, Overridingness, or Just Normativity?

In discussion of the authority of morality, the three most common strains of thought have to do with its inescapability, its overridingness, and its normativity. I will

give a brief gloss of each of these senses of authority and how they apply in arguments against realism.

1.2.1 Inescapability

The inescapability of the moral is the idea that one cannot “opt out” of the demands of morality in the way that one can opt out of other kinds of norms, such as the rules of a game or the laws of a nation, by giving up or leaving.

Philippa Foot offers several different ways of understanding the idea that moral demands are inescapable, focusing on Kant’s idea that moral demands are categorical rather than hypothetical imperatives (Foot 1972).

Modern philosophers follow Kant in talking, for example, about the "unconditional requirement" expressed in moral judgments. These tell us what we have to do whatever our interests or desires, and by their inescapability they are distinguished from hypothetical imperatives (Foot 1972 p. 308).

She considers several different ways this claim can be understood and finds all of them to be problematic. One way to understand a categorical imperative is as a demand that it is binding on you independent of your individual desires. But Foot argues that the rules of etiquette work the same way; they are rules that are binding whatever desires you might have. For example, one ought to write a thank you note when one receives a gift. This is not an imperative that does not apply to someone simply because the person “has his own good reasons for ignoring this piece of nonsense, or who simply does not care about what, from the point of view of etiquette, he should do” (Foot 1972 p. 308). Such a person is still bound by the norms of etiquette, despite their views about their arbitrariness and so on. This person who fails to write a thank you note is rude,

regardless of whether or not they “choose to play the etiquette game.” So this cannot be the special kind of authority that morality alone possesses.

A second way of understanding the categorical nature of moral demands according to Foot, is that they “necessarily give reasons for acting to any man ...all ‘should’ or ‘ought’ statements give reasons for acting”(Foot 1972 p. 309). On this view what the unique authority of morality consists in is that it is a kind of irrationality to fail to take moral reasons to be categorical in the sense that Kant and Kantians seem to require.

However this is an extremely difficult position to defend. Foot argues that failing to act morally need not be either inconsistent or irrational. A person can only be irrational if he “...in some way defeats his own purposes, doing what is calculated to be disadvantageous or to frustrate his ends”(Foot 1972 p. 310). Here Foot is assuming what some have called the internal reasons proposal (in contrast with external reasons). The idea is that the only reasons we have to act on come from the inside rather than the outside, in that in order to have a reason to do something the act in question has to serve the satisfaction of a preexisting desire that one has.⁶⁹ Thus on this view there is a tight connection between motivation to act and having a reason to act. So according to a simplified version of internal reasons, it wouldn’t be irrational for someone to do something immoral, as long as they were acting in a way that is consistent with their desires and goals.

Even if we accept Foot’s arguments thus far, a feeling may remain that the requirements of morality are more stringent than the requirements of a system of rules

⁶⁹ Here Foot has to argue that there cannot be external, rather only internal reasons. For more on this idea see: (Williams 1980).

like etiquette. However, Foot speculates that this probably has to do with the way we are raised and the emphasis that is put culturally on following moral rules over other types of rules (Foot 1972 p. 312). Even if we abandon the idea of categorical imperatives, Foot's alternative is that people have other reasons for acting morally, that have to do with the value they place on other people's well-being, for example (Foot 1972 p. 312). The ends and reasons that people have to act morally may be contingent, but this is the only alternative to pretending that morality has the "magic force" of the categorical imperative (Foot 1972 p. 315).

In the context of a discussion on the authority of morality, whether or not people act rationally when they disregard moral considerations or intentionally act contrary to moral considerations is somewhat of a red herring. If it is the requirements of morality *themselves* that have the authority, then whether or not we act rationally or irrationally when we ignore them seems to be beside the point. Unless we think that the only kind of authority that morality could have is authority within a particular theory of rationality.

1.2.2 Overridingness

The overridingness of moral considerations is closely related to, but distinct from the idea that they are inescapable, especially in so far as the focus is on reasons and rationality. This is the idea that "moral verdicts are always in some sense supreme whenever they come into conflict with the verdicts of a distinct normative domain" (Gert 2013 p. 3).

Overridingness can be understood in several different ways as well, but is usually understood as having to do with reasons. Moral reasons trump other reasons for acting,

the classic examples being self-interested reasons. This might mean that it is irrational to prioritize another kind of reason over a moral reason. But the claim does not have to be understood this way. It can also be understood as saying that moral considerations should always have priority over other kinds of concerns.

Kramer puts this kind of overridingness in terms of the higher degree of normative strength, force, “unremittingness,” “insistence,” or stringency, that moral obligations have when they conflict with other kinds of considerations, such as “prudence or aesthetics or supererogatory ethics” (Kramer 2009 p. 143). On this view, moral duties almost always override other kinds of normative requirements (Kramer 2009 p. 144).

Phillip Pettit explains that what separates moral considerations from other kinds of “grounds of assessment” is that morality has an “authoritative standing,” “priority,” or overridingness (Pettit 2001 p. 234-36). For certain kinds of values, for example the values proscribed by the law, an agent can justify ignoring them by appeal to moral considerations. But this kind of justification does not run in both directions.

What distinguish ethical justification is not the particular norms it invokes but rather the role it plays in relation to justification of other sorts. We will say that the criterion of ethical as distinct from other kinds of assessment is precisely the fact that it enjoys a certain sort of priority in people’s justificatory practices (Pettit 2001 p. 235).

In other words, an agent “cannot expect to be able to justify indifference to such a criticism [a moral one] by arguing that it would be wrong in terms of etiquette or prudence or law, or indeed in any other terms, to honor the demands of ethics” (Pettit 2001 p. 234-36).

A related formulation of overridingness says that: “the totality of practical reasons, taken together, always uniquely favors acting in line with a moral verdict, even though it may not always be irrational to act in ways that are not uniquely favored in this way” (Gert 2013 p. 1). The basic idea here is twofold; first, it could be that moral duties are overriding because rationality, taken as a whole, considers all kinds of reasons, moral, prudential, aesthetic, etc., and it happens that moral requirements also happen to be the most rational. The second strain here is that moral reasons already take into account other kinds of reasons, such as prudential reasons. So the fact that a particular action is in one’s own self-interest has already been taken into account if a moral requirement tells us to act against it (Kramer 2009 p. 144).

Foot also uses the term “overriding” to describe one of the purported features of morality (Foot 1978). Foot attacks the notion she attributes to D.Z. Phillips (1977) that “to anyone who cares about morality moral considerations must be the most important of all considerations” (Foot 1978 p. 181). Foot points out that people act contrary to their moral judgments all the time without being bothered by it (Foot 1978 p. 185-6). She takes this to show that most people take moral considerations to be merely one set of considerations among several competing sets, that also include etiquette and self-interest, for example (Foot 1978 p. 184). Foot explains that while it might be morally required to prevent one’s party guests from driving home drunk, etiquette requires that one should neither restrict one’s guests’ level of alcohol consumption or embarrass them by suggesting that they shouldn’t drive (Foot 1978 p. 184).

The sense we have that moral considerations are overriding, she concludes is again, as in her analysis of the categorical nature of moral demands, simply a result of

the way morality is taught. Whereas etiquette is explained to children “a rigid set of rules that are on occasion to be broken;” “moral rules are not taught as rigid rules that it is sometimes right to ignore; rather we teach that it is sometimes morally permissible to tell lies (social lies)...” (Foot 1978 p. 186).

In this debate between Philips and Foot, questions of the authority (in this case the overridingness) of morality and questions of moral motivation are being conflated. It seems terribly obvious that people who really care about morality act in immoral ways routinely and fail to feel remorse or guilt. But it is difficult to see how this psychological observation can give us insight about the nature of the moral demands or requirements themselves. On a more charitable interpretation of Philips’ position, moral considerations *should* override other considerations in all agents. If he were to put his point this way then he would clearly be addressing the authority of morality and not motivation. In that case Foot’s response misses the mark.

1.2.3 Normativity

Although the chapter is on the authority of morality, both of the objectors I consider, Korsgaard and Parfit, argue that realists and naturalist realists respectively, have a problem explaining normativity. Normativity is a broader concept than the inescapability, categoricity, or the overridingness of moral requirements. Normativity is arguably not only a feature of ethical judgments, facts, and properties, but is also often invoked in aesthetics, law, epistemology, philosophy of mind, semantics, rationality, and even logic. Normativity is about the way that certain theories, concepts, or facts serve as standards or rules for evaluating character. Norms are rules or standards that are

meant to serve as guides to decision making and action. This is because normativity has to do with what we have reason to believe, do, or feel. I will briefly explain why I think I am entitled to use objections that claim to be about the normativity of morality in this way.

Korsgaard's objection to realism has two important features that keep it close enough to the authority objection to make my case. First, part of her charge against the realist is that he has no response to the moral skeptic who asks "Why should I be moral?" This sounds like a plea for the authority of morality to be justified. Second, it seems she sees the fact that her view can capture in a certain sense, why morality is inescapable, to be a virtue of her position. The account that she gives has to do with features of our agency.

For Parfit, one of the core problems with non-analytical naturalism is that it essentially eliminates normativity by claiming that identities exist between natural facts and normative facts. Parfit's conception of normativity is essentially tied to reasons.⁷⁰ What is important to figure out about Parfit's account is why he thinks the elimination of normativity is a loss. Although it is not obvious, there is a way in which Parfit's worry about the disappearance of normativity in non-analytical naturalism is close to Korsgaard's charge that the realist has no response to the moral skeptic.⁷¹ Parfit writes, "...normativity is best understood as involving reasons or apparent reasons. Things matter only if we have reasons to care about them" (Parfit 2011 p. 269). Reasons are

⁷⁰ The dialogue between Parfit and Korsgaard on their respective views is a fascinating one but would take me too far afield to be explored here.

⁷¹ Parfit would probably deny this, since he thinks that Korsgaard is exclusively focused on the motivational sense of normativity. I think Korsgaard slips into that way of thinking sometimes, but her account is not reducible to motivation.

“where the buck stops when we are asking what really matters - or seeking decisive guidance as to what we ought to think and do” (Railton 2013 p. 2).

1.3 Distinguishing Authority from Motivation

Distinguishing the authority objection from the internalism objection is a way of responding to premise 1 of the authority objection, the claim that morality has a certain kind of authority over us. Making this distinction has been one of the continuing and central themes of the essay thus far. If we discover that what an objector to naturalist moral realism really means by authority is that moral judgments (or facts-more implausibly) necessarily motivate then this is really just the internalism objection all over again, which has been dealt with in Chapters 2 and 3.

A conflation of the two objections is relatively common in the history of moral philosophy. For example, John Stuart Mill, addressing the following list of objections to his theory, provides an example of the conflation of the internalism objection and the authority objection:

The question is often asked, and properly so, in regard to any supposed moral standard - What is its sanction? What are the motives to obey? Or, more specifically, what is the source of its obligation? When does it derive its binding force? (Mill 1863 p. 38).

Here, Mill’s naturalist theory of morality is confronted with both the internalism objection and the authority objection. The internalism objection is the question posed to Mill about how his theory of morality can motivate. While the authority objection, is the question about the “binding force” of his moral theory. Mill however, does not differentiate these two challenges to this theory.

The way that Hare interprets the authority (or overridingness) of morality has to do with our psychology and slips into talk of motivation. Hare argues that "...[overridingness] is the thesis that one's morality is something that yields prescriptions which, as a matter of psychological fact, one lets override all other prescriptions. Hare argues that 'critical moral principles', that is, judgments about what ought morally to be done, all things considered, are 'overriding' in this sense" (Hare 1981 p. 24, 55-62; Copp 1995 p. 77).

A similar conflation of motivation and authority can be found in David McNaughton's criticism of externalist moral realism (McNaughton 1988). McNaughton stipulates that for the externalist moral realist what makes an action right or wrong is independent of the agent's desires. However, for the externalist moral realist, he charges, whether or not a moral judgment has any authority over an agent is dependent solely on the agent's desires. According to McNaughton's criticism, the externalist moral realist must say that while an agent may "recognize" that "inflicting unnecessary suffering on others is morally wrong" this judgment gives them no reason not to inflict that suffering, unless they also have a desire not to do so:

If he lacks the appropriate desires then he has no reason to act in accordance with moral requirements-they are not seen as authoritative...Externalist realism is happy to view moral questions as factual ones, but distances them from motivation in its claim that moral commitments do not, in themselves, provide the agent with reason to act...Thus both sides must claim that the apparent authority of moral demands is an illusion (McNaughton 1988 p. 48-9).

He charges that the externalist moral realist cannot account for an important aspect of moral phenomenology, the experience of moral authority. But here he assumes that if an agent does not happen to have the desire to act morally then morality is not authoritative.

Another example of the conflation of the internalism objection and the authority objection to naturalist moral realism is found in Mackie. Mackie argues that:

On a naturalist's [realists] analysis, moral judgments can be practical, but their practicality is wholly relative to desires or possible satisfactions of the person or persons whose actions are to be guided (Mackie 1977 p. 33).

Mackie is saying that naturalists usually embrace motivational judgment externalism, which is the view that when one makes a moral judgment one is not necessarily or inherently motivated to act on that moral judgment. Instead, on motivational judgment externalism, moral judgments are only contingently connected with motivation to act morally (See Chapter 2). Mackie claims that on a naturalist realist view, the action guiding force of morality is tied to the contingent desires of human beings. This can be understood as a comment about moral motivation or a comment on the authority of morality. Understood as a comment on moral motivation, it is simply a statement of motivational judgment externalism. If this is the right way to understand what Mackie is saying then he is correct; externalism does claim that moral judgments are only contingently motivating. However, as I discussed in Chapter 2, this view is actually superior to the competing view, internalism.

A second way to understand what Mackie is getting at in this passage has to do with the authority of morality. Moral claims are thought to make authoritative demands on our conduct. This may be the kind of "practicality" to which Mackie is referring. Mackie may be concerned that the naturalist ties the authority, the demandingness or authority of morality, to the contingent desires of human beings. Kant articulates this point when he says that moral laws are categorical rather than hypothetical.⁷² However,

⁷² Working out exactly how Kant's claim about categorical imperatives relates to the authority of morality is taken up by Philippa Foot in "Morality as a System of Hypothetical Imperatives" (Foot 1972).

Kant's observation does not mean that moral laws are inherently tied to motivation; it means that we are obligated to act according to the moral law regardless of our desires or goals.

More recent conflation of authority and motivation objections can be found in Copp and in Seriol Morgan who takes himself to be strengthening Copp's point (Morgan 2006). Copp claims that normativity is "internal" to moral properties. This means that moral properties are necessarily normative. Copp argues that the naturalist moral realism of Boyd, Brink, Railton, and Sturgeon make normativity external to moral properties, rather than "essential" to them (Copp 2007 p. 4). Copp goes on to cite as his proof that these realists make normativity external to moral properties, the fact that they make motivational force of moral judgments external to moral properties, which is a different matter (Copp 2007 p. 4).

According to Copp, the Cornell realists are not able to give an account of how a "natural property could be normative" (Copp 2007 p. 4). It is surprising that Copp makes this mistake; given the fact that elsewhere he distinguishes between internalism about motivation and internalism about normativity. In fact he embraces externalism about motivation, along with the Cornell realists (Copp 2007 p. 6).

Morgan argues that realism espouses a "queer" metaphysics of morals that is incompatible with the kind of special status that we take morality to have. In other words, there is a central feature of morality that naturalism cannot accommodate. He variously refers to this feature of morality as its "practicality," "normativity," "categorical nature," its status as giving us "universally compelling standards," and "reason giving" (Morgan 2006). Morgan's argument is that while naturalism might be able to tell us that

moral standards are objectively real, it has to concede that there are other sets of standards that are just as objectively real, like a Nietzschean set of values having to do with nobility and ruthlessness (Morgan 2006).

There are multiple sets of real values and we know they are real because they figure in the best explanations of things. However, according to Morgan, naturalism has no way of giving any one set of values or moral standards any special status or significance. This is because, according to Morgan's argument, naturalists are externalists about moral motivation; "on the naturalist account of normativity, an individual has reason to act morally if she has a motivational profile which contains appropriate sympathetic desires and feelings" (Morgan 2006 p. 333).

While it may be true that many naturalists are also externalists about moral motivation, there is no reason to think that they have to identify the authority of morality with contingent features of human beings, like our sentimental propensities. In addition to the conflation of moral motivation and morality's authority, Morgan assumes that to be a naturalist is to deny rationalism. This is not a point I focus on but it is significant to note that to be committed to naturalist moral realism one need not reject moral rationalism, understood as the view that if one has a moral obligation this means necessarily that one has a reason to act on it (Shafer-Landau 2012 p. 159).

Even if we assume that naturalists are committed to anti-rationalism, Morgan still is not entitled to attribute to the naturalist the claims that on their view "naturalism grounds the normativity of the properties that guide the actions of persons of a particular evaluative outlook in their contingent non-cognitive states" or that "normativity *springs*

from contingent sentiment, our sentiments differ from theirs, that is all there is to it”
(Morgan 2006 p. 337).

Morgan’s argument offers a perfect example of the kinds of criticisms of naturalism or realism that come from merging together these two distinct issues: how do our moral judgments motivate (cause) us to act and why our moral judgments *should* motivate us to act (the authority of morality). Realism, as I have been defending it, has two central components, a metaphysical component, about the ontological status of morality and a psychological component, having to do with the nature of our moral thoughts. The psychological question of whether or not we are motivated to act on our moral judgments cannot even tell us about whether or not we have reasons to act on our moral judgments (unless we assume reasons internalism). So it certainly cannot definitively answer questions about the authoritativeness or morality.

The internalism and authority objections should be differentiated because whether and how morality motivates us to act does not bear on the place that morality should have in our lives and decision-making. We may have no motivation to do things that we should do, and we may be extremely motivated to do things we should not do. The most common way of making this point is to say that motivation has to do with whether or not we care about moral facts and properties, whereas, the authority of morality deals with whether or not we have reason to care or a justification for caring about the moral facts and properties.

2. The Authority Objection: Korsgaard’s Challenge

Korsgaard is a good example of a comprehensive rationalist neo-Kantian account of the authority of morality. At the same time, her view is close to moral realism in that she is a cognitivist and takes morality to be objective. Korsgaard explicitly challenges moral realism on the grounds that it cannot account for the authority of morality, what she calls the normativity of morality.

As part of her overall defense of Kantian constructivism, she presents a version of the authority objection against moral realism, arguing that her view can explain the authority of morality, while realism cannot. She argues for premise 1 of the authority objection, that morality has a certain kind of authority over us; she calls this the normativity of morality. She also argues for a version of premise 2. Recall that premise 2 says that moral realism precludes this kind of authority. She argues that moral realism cannot provide an explanation of the normativity of morality. Korsgaard argues that realism cannot give an account of normativity (authority), but her view, Kantian constructivism can, and is therefore superior.

Unlike Mill, McNaughton, and Mackie, Korsgaard recognizes formally the distinctiveness of the question of the authority of morality calling it the “normative or justificatory” problem, and elsewhere the normative question (Korsgaard 1996 p. 13). An answer to the normative question, according to Korsgaard, must explain why we *should* care about morality, not just why *do* care about morality (Korsgaard 1996 p. 23).

The normative question is a first-person question that arises for the moral agent who must actually do what morality says. When you want to know what a philosopher’s theory of normativity is, you must place yourself in the position of an *agent* on whom morality is making a difficult claim. You then ask the philosopher: must I really do this? Why must I do this? And his answer is his answer to the normative question (Korsgaard 1996 p. 16).

And later she writes:

If someone falls into doubt about whether obligations really exist, it doesn't help to say, "Ah, but indeed they do. They are real things." To see this, imagine a case where morality requires you to face death rather than do a certain action. You ask the normative question: you want to know whether this terrible claim on you is justified. Is it really true that this is what you must do? (Korsgaard 1996 p. 38).

Korsgaard charges that the realist does not, and cannot, provide an adequate response to the normative question. The realist's difficulty in answering the normative question comes from moral realism itself, whether naturalist or not.⁷³

Why is it that Korsgaard thinks that realism has a harder time than constructivism (or any other metaethical position for that matter) answering the normative question? Korsgaard rightly points out that realism is a metaphysical position. The type of realism that Korsgaard is addressing is what she calls "substantive realism" as opposed to procedural realism. Substantive realism:

...is the view that there are answers to moral questions; that is, that there are right and wrong ways to answer them...that there are answers to moral questions *because* there are moral facts or truths, which those questions ask *about* (Korsgaard 1996 p. 35).

The substantive realist "... thinks that there are correct procedures for answering moral questions because there are moral truths or facts that exist independently of those procedures... the procedures for answering normative questions as ways of finding out about a certain part of the world, the normative part" (Korsgaard 1996 p. 36). Korsgaard identifies herself as a procedural moral realist, which is the position that, "there are answers to moral questions...there are right and wrong ways to answer them...But procedural realism does not require the existence of intrinsically normative entities"

⁷³ Other forms of the authority objection apply strictly to naturalist forms of moral realism See: (FitzPatrick 2008).

(Korsgaard 1996 p. 35) (From this point forward in the Korsgaard discussion I will refer to 'substantive realism' as simply 'realism').

The realist conception of the task of moral philosophy is part of their problem, she argues. Since realists see themselves as discovering and documenting the normative part of the world, it is hard for them to be able to show how those discoveries, the moral reality, moral facts, and moral properties, connect up with human action and practical decision making in the right way. Korsgaard says that realists see ethics as a branch of applied knowledge and this is the source of the problem.

In contrast, as a constructivist, Korsgaard sees ethics as providing solutions to practical problems:

...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. The role of the concept of the right, say, is to guide action... (Korsgaard 2003 p. 116).

Initially, there seem to be two options for locating the source of normativity, or answering the normative question, according to Korsgaard:

So we are faced with a dilemma. If we try to derive the authority of morality from some natural source of power, it will evaporate in our hands. If we try to derive it from some supposedly normative consideration, such as gratitude or contract, we must in turn explain why that consideration is normative, or where its authority comes from. Either its authority comes from morality, in which case we have argued in a circle, or it comes from something else, in which case the question arises again, and we are faced with an infinite regress (Korsgaard 1996 p. 30).

A divine command theorist or a social contract theorist faces the same objection according to Korsgaard. They have not given a justification for the obligation to obey the divine commands or the obligation to honor the social contract. If the source of the authority of morality is located in our desires or some set of emotions, then morality only

seems to apply to us if we have those particular desires or emotions. And so it loses the very authority, categoricity, or normativity that needs to be explained.

Korsgaard concludes that neither of these responses is adequate. The realist refuses to answer or even address the normative question. Instead of telling the moral skeptic why he is obligated to be moral, the realist, “dig[s] in his heels” and says that “[o]bligation is simply there, part of the nature of things (Korsgaard 1996 p. 30).

According to the realist, the authority or normativity of morality is something that cannot be further explained (Korsgaard 1996 p. 33). In other words, the realist’s response to the normative question is to dismiss it altogether (Korsgaard 1996 p. 30-32). Korsgaard finds this response to moral skepticism in the work of Richard Price, Samuel Clarke, H.A. Pritchard, W.D. Ross, and G.E. Moore, who, she explains simply assert that right actions are “intrinsically right” or have their fit-to-be-doneness built into them (Korsgaard 1996 p. 30-32).

As a metaphysical position, Korsgaard argues that realism falls victim to the same objection that divine command theory is felled by, which she calls a Kantian objection. Just as divine command theory is not equipped to respond to the question, ‘why am I obligated to do what God commands?’ realism is not equipped to answer the question ‘why should moral facts or properties play any role in my practical decisions about what to do and how to live my life?’

Another way of putting the problem is that Korsgaard sees a gap for the realist, between knowledge of moral facts or properties and the application of that knowledge to deliberation and action. Fitzpatrick explains the problem:

...in order to apply such a piece of knowledge and come to be motivated, Korsgaard argues, one would need another norm dictating, for example that one

ought (or is obligated) to perform good actions. Where, then is that norm supposed to come from? It cannot itself be derived from the above knowledge of the goodness of an action, for the very question is why one should care about such goodness and apply such knowledge (FitzPatrick 2005 p. 656-657).

Any moral facts that the realist argues exist require "...some further norm prescribing the application of such knowledge; and that further norm cannot itself be understood on the realist model, as a known normative truth to be applied, without just raising the same problem again"(FitzPatrick 2005 p. 656-657). This is why Korsgaard thinks that realism cannot answer the normative question. Korsgaard takes this to be a variation on the open question argument: "If it is just a fact that a certain action would be good, a fact that you might or might not apply in deliberation, then it seems to be an open question whether you should apply it" (Korsgaard 2003 p. 112).

This creates an infinite regress for the realist, whose only answer can be to defend the existence of *another* moral fact that says that we ought to apply the original moral facts to our deliberations, but why should we apply this new moral fact to our deliberations?

For even if we know what makes an action good, so long as that is just a piece of knowledge, that knowledge has to be applied in action by way of another sort of norm of action, something like an obligation to do those actions which we know to be good. And there is no way to derive such an obligation from a piece of knowledge that a certain action is good. A utilitarian thinks an action is good because it maximizes good consequences and a virtue theorist thinks it is good because it is kind of brave, But how is it supposed to follow that it is to be done? (Korsgaard 2003 p. 111).

In order to avoid circularity, the realists might argue that "some things are intrinsically normative"(Korsgaard 1996 p. 33). The intrinsically normative things are just the things that we have to do. The realist takes up the following strategy:

Having discovered that obligation cannot exist unless there are actions which it is necessary to do, the realist concludes that there are such actions, and that they

are the very ones we have always thought were necessary, the traditional moral duties... (Korsgaard 1996 p. 34).

But this realist retreat to intrinsic, necessary, normativity cannot answer the normative question, Korsgaard argues. Especially when we imagine the first person position of an agent whose moral obligations come into sharp conflict with their own self-interest. The agent takes on the role of the moral skeptic and asks whether she is really obligated to do as morality requires.

2.1 Responding to Korsgaard's Challenge

Moral realism needs to have a strong response to Korsgaard's challenge and the normative question. I respond to Korsgaard's criticism in four related ways. First, I think that she mischaracterizes realism's resources for answering the normative question. That is, she assumes that all the realist can do is stamp her foot and say that the normativity is just there. I think realism can do more than that without falling into the vicious circularity that she suggests it will. Namely, realism can give answers that come from first order normative theories about why we ought to do what is morally required. If this is not enough and the imagined situation of the moral skeptic still has force, it might be that there is a subtle conflation of justification and motivation going on. I expect that Korsgaard would not be satisfied by that response that realism can rely on first order normative theories to answer the normative question that she is asking.

I offer a second, Prichard-style response to Korsgaard's challenge. Prichard famously argued that the question 'why should I do what I am morally obligated to do?' is a confused question. It is self-evident that one ought to do what one is obligated to do. On his analysis any attempt to answer the question that Korsgaard poses will either

be trivial (it will give a reason to do one's duty in moral terms) or it will change the subject by giving an answer in nonmoral terms.

Third, I argue that Korsgaard's alternative to realism, Kantian constructivism, cannot adequately answer the challenge she poses. Her own view faces the same regress with which she charges the realism.

Fourth, I suggest that if all metaethical theories have difficulty answering. Korsgaard's strenuous formulation of the normativity question, realism fares better than many others on the grounds that it captures objectivity in a unique way and that objectivity is ultimately a large part of what we are talking about when we talk about the authority of morality (See Section 5).

2.2 Is Korsgaard Conflating Justification and Motivation?

The obvious response to the moral skeptic who asks whether he is really obligated to do what morality requires of him, is to lay out the reasons why the particular course of action is indeed morally required. What those reasons are and how persuasive they are is a matter of first order ethical theory. Depending on the first order ethical theory that the person answering the normative question is working with, they may appeal to the suffering of others affected by the agent's actions, the respect that we are all required to show for fellow rational beings, the rights of other people, etc.

Korsgaard recognizes that first order moral theory can give us straightforward answers to this question; she writes, "a utilitarian might say that the rule says that certain kinds of actions are to be performed (the principle of utility)"(Korsgaard 2003 p. 111). However, she thinks this is not enough to answer the normative question, for two

reasons. First, a theory like utilitarianism does not explain “why there is such a rule” (the rule of utility). Second, utilitarianism does not explain “...why we should conform to the rule,” (Korsgaard 2003 p. 111). For Korsgaard, what needs to be further explained is why the rule has any normative force itself or whether any proposed rule is obligatory to follow.

Korsgaard draws a strong contrast between substantive realism and her constructivism. Part of this contrast has to do with the way that realists take themselves to be “describing normative reality,” while constructivists take morality to be “solving practical problems of what to do in conflict” (Hussain and Shah 2006 p. 289). But this is unfair to the realist because not only will they agree that the purpose of morality is “about finding the solution to the problem of what to do,” they argue that “[c]orrectly describing normative reality, discovering which actions have the property of to-be-doneness, answers the question of what one should do” (Hussain and Shah 2006 p. 289).

In the test case that Korsgaard uses for answering the normative question, doing the right thing requires great personal sacrifice on the part of the agent in question. Of course simply telling the agent that it is a *real* moral requirement will not help. Facing death in order to fulfill a moral requirement is an extreme situation. Citing moral reasons drawn from either a philosophical ethical theory or common morality probably will not help either. When imagining this conversation, convincing someone on a first person level that they are obligated to risk their life, it is very hard to differentiate questions of justification from questions of motivation. It is easy to imagine the agent saying, yes I

see why risking my life is the morally required course of action, but then hesitating and asking for further justification.

But I wonder at that point in the conversation if the skeptic is simply expressing the fact that they just cannot gather up any *motivation* to do what they are morally required to do. Much of the force of considering this issue in a first person context is that it is easy to imagine oneself struggling to accept, not only that morality requires a particular action in challenging circumstances, but further that moral obligations really are obligatory in general, if they are the kinds of things that require extreme self-sacrifice. In these cases one has an overwhelming conflicting motivation to take a different course of action. Korsgaard points out that, the realist may be confident in the obligatory nature of moral requirements, but not everyone is (Korsgaard 1996 p. 39-40).

In addition to the first person perspective, our interpretation of the skeptical question may also be influenced by the fact that a moral requirement is being posed that involves serious self sacrifice. What if the example was much more mundane and involved minimal personal sacrifice in order to fulfill a moral requirement. Imagine morality requires a young and healthy agent forfeit their seat on the bus for an elderly person with a broken leg. The agent has only two stops left before they get off the bus and so surrendering their seat to the elderly person will create only a little discomfort for the agent. In this case, our agent is convinced that giving up the seat is the right thing to do (on whatever first order ethical theory we are assuming is correct). Now he asks, "Am I really obligated to do what is morally required? Must I make this sacrifice?" When the level of personal sacrifice involved in fulfilling one's moral duty diminishes, so does the rhetorical force of the question. Although Korsgaard explicitly distinguishes

between motivation and normativity, part of the force of this question seems to be coming from the motivational issue. This leads to the next response, which is the Prichard-style response. If indeed, the person is asking am I “obligated to do what I am obligated to do?” then the answer, rightfully, can only be ‘yes.’

2.3 Prichard-Style Response to Korsgaard⁷⁴

The mistake that moral philosophy famously rests on, according to Prichard, is trying to provide a foundation for our moral obligations outside of duty itself. For the Prichard-style response, the problem is not that any answer to the first person demand for an explanation of the normativity of moral obligations will not satisfy the *agent* (that criticism will be addressed in section 2.4 on why Korsgaard’s answer the moral skeptic is unsatisfactory). The problem is that the question itself is conceptually confused and self-defeating (Stern 2010 p. 1). The answer to the skeptic’s question will have to place the source of the normativity of morality outside of morality itself. The Prichard style response says that:

...to take skepticism seriously in the way that Korsgaard does, is to assume that morality needs some extra-moral basis; however, to be moral is precisely to think the moral reasons one has to act are compelling in themselves without any such basis for them being required by someone who is a genuine moral agent (Stern 2010 p. 3).

Prichard argues that when the question is asked “why should I do these things [which I am morally obligated to do]?”(Prichard 1912 p. 22) There are two kinds of responses that are available to answer this question: “the reason ‘why’ is stated in terms either of

⁷⁴ Another historically grounded way to respond to Korsgaard is by making Foot’s point that just because someone is immoral does not necessarily mean they are irrational.

the agent's happiness or of the goodness of something involved in the action”(Prichard 1912 p. 22). Of course Prichard argues that neither of these answers is suitable for the following reasons.

The answer posed in terms of “the agent’s happiness” is essentially an egoist answer; this is the type of answer Prichard takes Plato, Aristotle, Butler, Hutcheson, Paley, and Mill to have given (Prichard 1912 p. 22). The problem with the egoist answer: “[t]he answer is, of course, not an answer, for it fails to convince us that we ought to keep our engagements; even if successful on its own lines, it only makes us want to keep them” (Prichard 1912 p. 23). In other words, that a particular action is in our own self-interest or will make us happy is “the wrong sort of reason for a person to conform to specific moral demands” (Gaus 2011 p. 57).

The answer given in terms of the goodness of the action might be either a consequentialist answer that says that one ought to do what will produce the most good or a Kantian answer that says that one ought to do what is good in itself. Prichard has two objections to the answer focused on either the instrumental or inherent goodness of the action. Prichard understands the instrumental goodness of the action to mean that the action creates a good state of affairs. This answer, he argues, assumes “that what is good ought to be” (Prichard 1912 p. 24). And this is not necessarily the case.⁷⁵ He argues that this does not fit with our ordinary thinking about right and wrong. When we think about why we ought to pay back a debt that we owe someone, we do not usually think that this is the action that will create a state of affairs that is most good (Prichard 1912 p. 25).

⁷⁵ Prichard gives another objection to the utilitarian answer to the question which has to do with the idea that an ought can only be derived from another ought. See (Prichard, 1912 p. 24).

The problem with thinking that the inherent goodness of the action can provide an answer to the question, why ought I to do my duty, is that it involves an incorrect assumption about what makes an action good. According to Prichard, only the motivation of the agent makes an action good (Prichard 1912 p. 26). However the rightness or wrongness of an action is independent of the agent's motivation (Prichard 1912 p. 26). It is often the case that we do the right thing without a good motivation, as in the case when we pay our bills not out of our sense of obligation to repay our debts, but out of fear of being sent to jail (Prichard 1912 p. 27).

2.3.1 Korsgaard's Response to Prichard

Korsgaard is well aware of the fact that she is committing what Prichard would call moral philosophy's mistake, by asking the normative question in the first place. Korsgaard addresses Prichard's argument and includes him in her list of realists who refuse to answer the normative question:

Prichard argues that it makes no sense to ask why you should be moral. If I give you a moral reason - such as, "it is your duty" - then my answer is circular, since it assumes you should be moral. If I give you a self-interested reason - such as, "it will make you happy"- then my answer is irrelevant. That is not the reason why you should be moral; you should be moral because it is your duty. If a question admits only of answers that are either circular or irrelevant then it must be a mistake to ask it. And if that is the question of moral philosophy, Prichard thinks, then moral philosophy rests on a mistake. Obligations just exist, and nobody needs to prove it (Korsgaard 2003 p. 32).

Prichard, like other realists, Korsgaard argues, falsely assume that stipulating the existence of inherently normative entities will end the discussion (Korsgaard 1996 p. 33). Korsgaard thinks she can do better than this by telling the skeptic that it is part of

the nature of being an agent, taking anything to be a reason for acting that one value the humanity of oneself and of others.

Briefly, Korsgaard's solution to the normative problem has to do with our nature as deliberative creatures, agents that act based on reasons. As such, we must form practical identities which guide our decision making. Korsgaard observes that we have a unique "capacity for self-conscious reflection about our actions." This structure is what creates "normative problems" for human beings. We are the types of beings that make choices and act based on reasons; we deliberate about what we should do (Korsgaard 2009). If we did not have a conception of our practical identity, we would not be able to formulate reasons for our actions.

Having a practical identity is not optional according to Korsgaard. A practical identity Korsgaard argues is "a description under which you value yourself and you find your life worth living and your actions to be worth undertaking" (Korsgaard 2009, p. 24). Although we have several practical identities, we all share the identity of being a human being (Korsgaard 1996 p. 212). We value this practical identity as it is the source of everything else we value. If we did not value this identity, she argues, we would not be able to value anything else. If Korsgaard can establish that we must value our own human identity, then, she argues, we are "rationally required" to value humanity "in the persons of others" as well (Korsgaard 1996 p. 121).

Every subject constructs their own practical identity and the moral law is the law of that identity. She writes:

The reflective structure of human consciousness requires that you identify yourself with some law or principle which will govern your choices. It requires you to be a law to yourself. And that is the source of normativity. So the argument

shows just what Kant said that it did: that our autonomy is the source of obligation”(Korsgaard 1996 p. 103-104). And this is what gives “us a kind of authority over ourselves, and it is this authority which gives normativity to moral claims” (Korsgaard 1996 p. 19).

2.3.2 Evaluation of Korsgaard’s Response to Prichard

Korsgaard’s response to Prichard is problematic in several ways. First, the way that Korsgaard interprets Kant, as a “radical constructivist,” leads her to have to say that the value of rational beings is not inherent in the beings themselves, as a realist reading of Kant would lead one to say (Watkins and FitzPatrick 2002 p. 360). This is problematic on two grounds, first it is inconsistent with some of our firmly held moral commitments. We don’t think that we should value people as rational beings because of the way that this connects to our ability to will. Second, it is problematic in terms of Korsgaard’s efforts to explain normativity in a way that is unavailable to realists. The normative force of the explanation that she can provide for morality seems far weaker than what the realist can provide. On her reading of Kant:

...the value of humanity ...is not simply a fundamental property of humanity, as a realist would have it, but winds up being further explicated in terms of a certain modal fact about the nature of willing: the alleged fact that we cannot get our wills to function unless we think certain things about ourselves (Watkins and FitzPatrick 2002 p. 360).

Watkins and Fitzpatrick worry that this view does not give the right explanations for certain moral facts and their normative force, like the fact that it is wrong to enslave a fellow human being. The reason that Korsgaard has to give, has to do with “...a story about what I have to do in order to exercise my will” (Watkins and FitzPatrick 2002 p. 361). But this does not seem right; people are not the sort of things that can be

property, possibly because of something having to do with their inherent value. It is their value that has “normative force for us” (Watkins and FitzPatrick 2002 p. 360-361). But instead of giving an account of this, Korsgaard’s view, gives a “... reduction of the value of persons to the commitments of deliberators arising from generic conditions on the exercise of rational agency” (Watkins and FitzPatrick 2002 p. 360-361).⁷⁶

Thomas Nagel makes a related criticism of Korsgaard, arguing that her version of constructivism is compatible with egoism, depending on “how we conceive of ourselves as reflective beings” (Nagel 1996 p. 203-204). Thus Korsgaard’s Kantian constructivism does not provide a stable or consistent basis for holding that humanity is intrinsically valuable, because “...the reflective standpoint can yield different results for different individuals because each person has his own reflective point of view” (Nagel 1996 p. 203-204).

One of the ways that Prichard considers that one might answer the question ‘Why should I do what I am obligated to do?’ would be in “terms ...of the agent’s happiness.” That is, one response would be that doing one’s moral duty aligns in some (or in a variety) of ways with one’s own self-interest. Korsgaard’s answer to the normative question is not obviously self-interested in any way. She is not saying that if you do not do your duty you won’t be happy or that you will be punished by a deity. However, on some readings of her work, she continues to fall victim to Prichard’s charge of asking an incoherent or meaningless question (Stern 2010, p. 4-5).

⁷⁶ This is restated when they write: “There seems to be a commitment at the meta-ethical level to a detour in the explication of another person’s value insofar as it is normatively significant for me through a certain modal understanding about me or about what I have to be committed to in order to exercise my will at all” (Watkins and FitzPatrick 2002 p. 362-363).

Both of these criticisms can be seen as Pritchard-style responses to Korsgaard, because they show that she attempts to provide an answer to the question why should I be moral from outside of morality itself.⁷⁷

2.4 Korsgaard Cannot Answer Her Own Normative Question

Korsgaard charges that the moral realist cannot and does not address the normative question or explain the sources of normativity. One way she presses this objection is by arguing that any acceptable answer to the normative question, must succeed in addressing someone who is in the first person position of the agent who demands a justification of the claims which morality makes upon him (Cohen 1996 p. 178).

In this section I will present three difficulties with Korsgaard's own proposed solution to the normative question. First, is a point made by G.A. Cohen about Korsgaard's demand for an answer to a first person request to the question of why ought I to do my duty? Second, are Fitzpatrick's and Copp's points that Korsgaard's solution to the normative problem, focused on valuing one's own humanity, cannot answer skeptical worries about the normativity of morality because people can consistently value other things in addition to their own humanity. A related problem comes up when we ask why an agent must value their own humanity in the way that Korsgaard insists they must. If this is a basic psychological fact about human beings, then it makes sense to raise the objection that Korsgaard's move from the inescapability

⁷⁷ Stern is the one that makes the connection to Pritchard.

of valuing something to the normativity of valuing something (namely humanity) is unsupported.

2.4.1 First Person Requests

The first person request for a reason why I am obligated to do what morality requires, the normative question, could be asked in either “the context of protest” or “the context of self-justification,” according to Cohen (Cohen 1996 p. 179). The context of protest occurs when doing the right thing is going to come at high personal cost for the agent, frustrating the agent’s other plans or desires (Cohen 1996 p. 179). Cohen wonders whether an “agent asking that question in so intransigent a spirit...could be satisfied by any theory Korsgaard’s included” (Cohen 1996 p. 178). Korsgaard’s answer will likely be insufficient; Cohen writes, “[i]t is useless to tell me that it lies in my practical identity to be thus obliged” (Cohen 1996 p. 180). Cohen explains that if a metaethical theory is required to answer the normative question in the context of protest it is likely that the agent’s “practical identity” has already been disrupted by the life circumstances that have led them to ask this question (Cohen 1996 p. 180). But none of this will be persuasive to someone who does not see why they are obligated to act morally. Telling them that it is inconsistent with their own sense of self or will does not apply to someone who does not feel the force of this (Cohen 1996 p. 182).

A person in a tight corner may not like the sacrifice a moral requirement demands of her, but she may reasonably be expected to recognize the normative force of that requirement for her if we have both justified the relevant principle, showing why it is true, and show how she as an agent has no choice but to adopt some practical principle in order to function as an agent and so must in effect choose between this true principle and other incompatible and thus false ones (FitzPatrick 2005 p. 689).

Although first person requests may be impossible to definitively answer, realism's unique claim to truth and objectivity may best position it to answer such a request.

2.4.2 Why should I value my humanity?

The broader objection here is that whatever reason Korsgaard gives for why we are obligated by moral obligations, her answer to the normative question, is open to the same kind of infinite regress that she accuses the realist position of being vulnerable to (Radzik 1999 p. 36; FitzPatrick 2005 p. 673).

...if it is a fact that I will be acting consistently with all the implications of my exercise of agency only if I make some sacrifice required by the formula of humanity, then it is so far an open question whether or not I should apply that fact in my deliberation and give it overriding weight. Why should I? Her own argument against realism may thus be turned around against her own account, again raising the prospect of an infinite regress of justifying norms (FitzPatrick 2005 p. 673).

The agent in the difficult situation can always ask about why he or she is obliged to value his or her own agency, above and to the exclusion of all else in this situation. And will then be faced with the same dilemma she accused competing accounts of value to face, having to locate the source of this value outside of morality (as divine command theory does) or insisting that the value is just there, in the nature

A related problem for Korsgaard's answer to the normative question is that while a rational person certainly values her own agency, she can simultaneously value other things. Sometimes acting morally will thwart her ability to achieve those other things she values (Copp 2007 p. 278). This does not mean that the person stops valuing his or her own agency. But they can ask, "[w]hy should I care about acting full consistently with all the implications of my exercise of agency...[a]nd why should I care about this

more than I care about anything else, especially when some onerous sacrifice may be involved?”(FitzPatrick 2005 p. 673).

Unless Korsgaard can say that it would actually be impossible for them to ask this question, then she has merely delayed the normative question. At times, she does come close to saying that it is impossible on her view for the normative question to be asked again in response to her answer to it: “we as agents should care about these issues of commitment and consistency because we have no choice: they are practically necessary for our very functioning as agents, which is our ‘plight’ as creatures with reflective consciousness”(FitzPatrick 2005 p. 673). Thus it seems that Korsgaard’s position involves the following unlikely fact:

...it is literally practically necessary in order for us to exercise agency at all (1) that we strive to act fully consistently with all the implications of the exercise of our agency (as revealed by Korsgaard’s and Kant’s sophisticated philosophical reflection) , and it is similarly necessary (2) that we care about this more than we care about all the things that might conflict with those implications, such as a central life project the pursuit of which might happen to conflict with the formula of humanity (FitzPatrick 2005 p. 673).

But obviously people act immorally all the time and so it cannot be that Korsgaard means that it is literally impossible for us to be inconsistent in this way (FitzPatrick 2005 p. 674). But it is not clear then what kind of impossibility she would be referring to.

2.4.3 From Inescapability to Normativity?

Another way of explaining the inadequacy of Korsgaard’s theory to deal with the authority of morality can be found by looking at her “move from inescapability to normativity” (Silverstein 2012 p. 4).

According to this ... theory, we should respond to skeptical doubts about the demands of morality not by digging in our heels and insisting as the realist does,

that the obligations really do exist, but rather by showing that those demands are inescapable (Silverstein 2012 p. 4).

According to Korsgaard, we do in fact value our own humanity and for reasons of consistency, are obligated to value the humanity of others. But as Silverstein points out, the fact that we value our own humanity does not tell us anything about whether or not humanity itself is actually valuable or not; [w]e might...be helplessly devoted to pursuing something which actually ought not to be pursued" (Silverstein 2012 p. 4-5). Valuing our humanity is an "attitude" that we "inescapably" hold, according to Korsgaard (Silverstein 2012 p. 6). But this matter is a contingent fact about us, human beings, but not about the nature of humanity itself.

Korsgaard argues that morality has a certain authority over us, she calls this normativity. And she argues that substantive realism (realism) cannot explain normativity, only her form of Kantian constructivism can. In this section I have argued that realism does not refuse to address or lack the resources to address the normative challenge that Korsgaard poses. But that the way that Korsgaard poses the challenge is too hard for any metaethical theory to answer. At the same time Korsgaard's constructivism faces its own difficulties in answering the normative challenge that she argues realism cannot address.

Korsgaard does not specify whether the version of realism she is referring to is naturalist or non-naturalist. The realism that I defend is naturalist. Thus it is important to consider a prominent critique of naturalist moral realism' account of normativity or the authority of morality that focuses specifically on the naturalist commitments of the theory. Parfit hold a non-naturalist realism and directs his argument against "non-analytical naturalism" (Parfit 2011).

3. Parfit's Normativity

As a nonnaturalist realist, Parfit takes normative properties to be irreducible non-natural properties. He claims that "normativity is essentially related to reasons"(Kieseewetter 2012 p. 466). This is the "reason-involving conception" of normativity. For Parfit, to deny the "irreducibility of normative facts" is to deny their existence altogether (Parfit 2011 p. 267). Normative facts are facts about what we have reasons to do; "[n]ormativity involves reasons or apparent reasons" (Parfit 2011 p. 268).

Reasons are, I believe, fundamental. Something matters only if we or others have some reason to care about this thing (Parfit 2011 p. 148).

What is normative are certain truths about what we have reasons to want, or will, or do (Parfit 2011 p. 424-25).

In this section I will give a brief account of Parfit's objections to naturalism. Parfit's objections fail for reasons particular to each argument. But there are also problems with his overarching approach. Parfit's objections to non-analytical naturalism seem to be far more applicable to analytical naturalism. Nor does Parfit distinguish between nonreductive non-analytical naturalism and reductive non-analytical naturalism. Second, Parfit's arguments use the same kind of intuitions that Moore relied on in the open question argument for the naturalistic fallacy.

3.1 Parfit's objections to non-analytical naturalism:

Despite Parfit's dissatisfaction with Korsgaard's constructivism, there are some striking similarities between Parfit's objections to naturalism and Korsgaard's objections to realism. Both worry about the possibility of moral skepticism in the sense that a

moral agent, even if he or she is aware of all of the facts of a situation, may always be able to ask, but why should I do what I am obligated to do? For Parfit, the problem in naturalism that gives rise to this worry is that fundamental reasons are omitted from a naturalist account. If naturalism is correct, then “we might be aware of pleasure and pain, weal and woe, hope and despair, helping and harming, truth and falsehood- everything that is of normative importance-yet absent recognition of independent reason relations, ‘that importance would be unknown to us-as it is unknown, for example, to some active intelligent cat’” (Railton 2013; Parfit 2011 p. 288). This means that we could be aware of everything that has normative *significance*, but if we do not recognize the fundamental reason giving force of those things that have normative significance, we have missed normativity all together (Railton 2013).

This echoes Korsgaard’s claim that realists have the wrong conception of what morality does; “goodness in action cannot just be a matter of applying our knowledge of the good—not even a matter of applying our knowledge of what makes action itself good” (Korsgaard 2003 p. 110). According to both views, having knowledge, or being aware of all of the facts that are of normative significance, and even knowing that an action, for example, is morally required, still leaves something significant out of the picture; it leaves out normativity.

Parfit recognizes that non-analytical naturalists are not trying to make claims about the meaning of moral terms; rather they are making claims about the nature of moral facts and moral properties. Parfit also appreciates that what makes such a view appealing is that it seems more consistent than other forms of realism with a naturalistic

world view that takes all facts to be natural facts.⁷⁸ Parfit takes non-analytical naturalism to say:

...we use some words to express concepts and make claims that are irreducibly normative, in the sense that these concepts and claims cannot be defined or restated in non-normative terms. When we turn to facts, however, there is no such deep distinction. All facts are natural, but some of these facts are also normative, since we can also state these facts by making irreducibly normative claims (Parfit 2011 p. 295).

Parfit presents three interrelated arguments against non-analytical naturalism, “the normativity objection,” “the triviality objection,” and “the fact-stating argument.”⁷⁹ The three arguments share some common features. All seem, to share a common heritage with Moore’s argument against the possibility of defining ‘good,’ especially in naturalistic terms, despite the fact that Parfit recognizes that naturalists are not talking about the meaning of moral terms. As part of the Moorean lineage of arguments against naturalism, Parfit’s objections are also open to some of the responses (in a modified form) given to the open question argument in Chapter 1. Parfit’s arguments rely on intuitions about what normative facts must be like. But it is hard to see why someone who does not already share Parfit’s nonnaturalist position and assumptions would have the same intuitions that he expresses.

Parfit, like Korsgaard, is explicit about distinguishing questions of normativity from questions of motivation. When discussing senses of “ought” that have to do with motivation or “rational necessity,” Parfit writes, “an irresistible impulse is not a normative reason. Nor is an impulse made rational by its ability to survive reflection on the facts.

⁷⁸ Here Parfit is characterizing the non-analytic naturalist’s position in terms of facts. I prefer to describe the position in terms of the properties that naturalists take to exist, but here I will use Parfit’s fact language.

⁷⁹ Several more arguments or sub arguments against moral naturalism can be identified in Parfit’s work. However, Parfit identifies these three arguments as his central objections to naturalism.

Even after carefully considering the facts, we might find ourselves irresistibly impelled to act in crazy ways” (Parfit 2011 p. 291). Ultimate normativity is not found in either motivation or in the rational requirement to behave morally.⁸⁰

Here I will present and respond to Parfit’s normativity objection and triviality objection. I leave out the fact-stating objection because he relies on the same intuitions about the incommensurability of the natural and the normative that inform the other two stronger arguments and so the third argument does not contribute anything significantly different to Parfit’s critique of naturalism. The fact-stating objection also assumes a controversial position about how the non-analytic realist understands the nature of identities between natural and normative properties (Copp 2013 p. 53). The fact-stating objection assumes that the non-analytical naturalist is committed to the “strong fact-identity thesis” which says that “if M is identical to N, then each fact to the effect that something is M is identical to the fact that that thing is N” (Copp 2013 p. 53). A discussion of this principle is not within the scope of this chapter.

3.2 Parfit’s normativity objection to non-analytical naturalism

Parfit takes non-analytical naturalism to be committed to the idea that “we make some irreducibly normative claims” but “there are no irreducibly normative facts” (Parfit 2011 p. 234). Parfit also thinks that it is committed to the view that true normative claims “state facts that could also be stated by making other, non-normative and naturalistic claims. These facts are both normative and natural” (Parfit 2011 p. 234).

⁸⁰ In the background here is a discussion about internal and external reasons; “To avoid confusion, we should use the phrase ‘a reason’ only in its external, irreducibly normative sense” (Parfit 2011 p. 290).

His argument against this view hangs on his ability to give evidence for the following set of claims. First, “normative and natural facts are in two quite different, non-overlapping categories” (Parfit 2011 p. 324). And second, that “natural facts could not be normative in the reason implying sense”(Parfit 2011 p. 324-325).

Some kinds of facts have been historically thought of as being “in a different category that physical facts,” like facts about vital force. But this thought turned out to be wrong; facts about vital force are not, in fact, in a *sui generis* category, conceptually different from natural facts. Parfit cites an ongoing debate about “whether conscious experiences could be the same as, or consist in, physical events in some brain” as an example of one of these questions about natural and non-natural facts that remains unsettled (Parfit 2011 p. 324).

Many kinds of thing, event, or fact are, however, undeniably in different categories. Rivers could not be sonnets, experiences could not be stones, and justice could not be— as some Pythagoreans were said to have believed— the number 4. To give some less extreme examples, it could not be a physical or legal fact that $7 \times 8 = 56$, nor could it be a legal or arithmetical fact that galaxies rotate, nor could it be a physical or arithmetical fact that perjury is a crime. It is similarly true, I believe, that when we have decisive reasons to act in some way, or we should or ought to act in this way, this fact could not be the same as, or consist in, some natural fact, such as some psychological or causal fact (Parfit 2011 p. 325).

The relevant conception of normativity here is the reason giving sense of normativity that Parfit defended earlier. Parfit’s basic point here is that normative facts and natural facts are simply two different kinds of facts that cannot be identified with each other. Even though it makes sense to say water is H₂O, “rivers cannot be sonnets” (Parfit 2011 p. 323).

So far this is not an argument, but merely a set of assertions that nonnaturalists make. Copp helpfully reconstructs Parfit’s normativity objection in the following way:

Any reductive naturalistic theory must identify each normative property M with some natural property N. ...any such property-identity thesis will *drain the normativity* from claims to the effect that things that are N are M. Intuitively, such claims are substantive and normative, but this is not so, if the naturalist is correct (Copp 2013 p. 37) [italics added].

Parfit clearly anticipates the non-analytic naturalist's response here, they are not analyzing the meanings of words, but rather suggesting the possibility of identity between facts or properties, analogous to cases of the identity between water and H₂O, or heat and molecular motion, which were discovered in an *a posteriori* way, that had nothing to do with the meaning of the words water or heat (Parfit 2011 p. 325).

However, Parfit argues that the analogy between to water and heat does not hold for the identity between natural and normative facts. In the case of water or heat the possible space of identities was "constrained by the relevant concepts" of water and heat (Parfit 2011 p. 325). In an earlier discussion of naturalism, he argues that, heat for example, has an "explicit gap that is waiting to be filled, since this concept refers to some property without telling us what this property is" (Parfit 2011 p. 302). The gap comes from the fact that we can identify the causes and effects of heat without really knowing what heat is. Whereas normative concepts like "wrong" do not have "explicit gaps that are waiting to be filled, in ways that would allow these concepts to refer to one or more natural properties" (Parfit 2011 p. 302). The concepts of water and heat "leave open various possibilities, between which we must decide on non-conceptual grounds" (Parfit 2011 p. 325). So the concepts of heat left it open whether it was molecular motion or phlogiston, it did not leave every possibility open.

But heat could not have turned out to be a shade of blue, or a medieval king. And if we claimed that rivers were sonnets, or that experiences were stones, we could not defend these claims by saying that they were not intended to be analytic, or conceptual truths. Others could rightly reply that, given the meaning of these

claims, they could not possibly be true. This, I believe, is the way in which, though much less obviously, Normative Naturalism could not be true. Natural facts could not be normative in the reason-implying sense (Parfit 2011 p. 325).

Something about the reason-implying concept of normativity prevents it from being identical to any set of natural facts. Parfit's suggestion that heat could not have turned out to be a king is surely intuitively plausible the way it is presented. As of yet, Parfit has not made the case that the naturalists identification of moral facts with natural facts is equally implausible.

Parfit's central way of illustrating the claim that there is a disanalogy between other non-analytic property identities and normative identities is with the following example. Imagine you are in a burning hotel, in order to save yourself you must jump into a canal:

Since your life is worth living, it is clear that (B) you ought to jump. This fact, some Naturalists claim, is the same as the fact that (C) jumping would do most to fulfill your present fully informed desires, or is what, if you deliberated in certain naturalistically describable ways, you would choose to do. Given the difference between the meanings of claims like (B) and (C), such claims could not, I believe, state the same fact. Suppose that you are in the top story of your hotel, and you are terrified of heights. You know that, unless you jump, you will soon be overcome by smoke. You might then believe, and tell yourself, that you have decisive reasons to jump, that you should, ought to, and must jump, and that if you don't jump you would be making a terrible mistake. If these normative beliefs were true, these truths could not possibly be the same as, or consist in, some merely natural fact, such as the causal and psychological facts stated by (C) (Parfit 2011 p. 326-7).⁸¹

The moral or normative fact here is (B) you ought to jump. The naturalist's fact is (C) If you deliberated in certain naturalistically describable ways; you would choose to jump (paraphrasing one of Parfit's examples of a naturalistic fact). Parfit's example of a naturalist's proposed fact that they would claim is the same as fact (B). Parfit claims that if (B) is true, it "cannot possibly be the same as, or consist in" (C). The obvious question

⁸¹ For a related, but distinct normativity objection to naturalism, see: (Dancy 2006).

to ask Parfit here is why this is the case. Especially, considering that naturalists would hope to give a much fuller account than the account given by (C). Parfit is ruling out the possibility that any kind of fact (C) that can be stated in non-moral language cannot be the same as fact (A).

3.3 Responding to Parfit's Normativity Objection

Copp and others have pointed out that here Parfit does not provide a strong reason to accept his claim that (B) and (C) cannot possibly be the same fact (Copp 2013 p. 47). Part of the impossibility of the identity between the facts stems from Parfit's claim that "reason-implying normative facts" are necessary facts, like the facts of mathematics (Parfit 2011 p. 326).

It makes sense to think that "...the concept of heat rules out the possibility that heat is a cabbage or a king" and "that the concept of rightness rules out the possibility that rightness is a rocket or a mountain lion or that it is the property of being a yellow rose" (Copp 2013 p. 47). Those would be very strange identities, but those are specific identities, not a whole category of types of identities, natural identities. So Parfit does not argue for the conclusion that moral facts or properties cannot possibly be natural facts or properties as a result of the concept of morality (Copp 2013 p. 47). Claiming that a concept rules something or another out requires more detail than Parfit provides. Without answering the questions in what way does the concept of morality rule out natural identities? or what is it about the concept that does so? All Parfit has given us is an assertion.

Parfit may answer this question by relying on his analysis of normativity as essentially tied to the reason giving nature of moral claims. However, this will not solve the problem. We would still demand "... to be given a reason to think that the concept of a reason ensures that the fact that a person has decisive reason to jump 'could not possibly be the same' as any natural fact'" as Parfit claims it could not be (Copp 2013 p. 47).

Here we might be reminded of Frankena's response to the naturalistic fallacy. Moore claims that while our moral theories can tell us about what kinds of things are good, they cannot find define good itself in terms of anything else. Thus good is simple and unanalyzable. For Parfit, normativity fundamentally has to do with reasons, which are non-natural and *sui generis*. As discussed in Chapter 1, there are many reasons to think that Moore's argument does not succeed as a devastating objection to naturalism.

Frankena pointed out that "the charge of committing the naturalistic fallacy" (though it is not a fallacy at all) "can be made, if at all, only as a conclusion from the discussion and not as an instrument of deciding it" (Frankena 1939 p. 465). The "generic" mistake being the naturalistic fallacy, the definist fallacy, is the mistake of "confusing or identifying two properties, of defining one property by another, or of substituting one property for another" (Frankena 1939 p. 471). But as Frankena shows, the definist fallacy is only a mistake if one has already established with independent evidence that the thing being defined (in Moore's case good) is indefinable (Frankena 1939 p. 473).

The same point applies to Parfit's normativity objection. Parfit claims that it simply cannot be that the natural fact that 'If you deliberated in certain naturalistically

describable ways, you would choose to jump out of the burning hotel' is the same as the normative fact that 'you ought to jump out of the burning hotel'. Parfit's only evidence for this claim is that there is an unbridgeable conceptual divide between these two sorts of facts.

Railton has pointed out that one of the central differences between rivers and sonnets, and natural facts and normative facts are the relations of constitution, supervenience, and determination. While it is true that "rivers do not supervene on sonnets...all the sonnets in the world do not determine which rivers there are...But all the natural facts of the world taken together do suffice to determine (metaphysically, not analytically) all the normative facts" (Railton 2013 p. 16). This means that if the moral facts were to change there would have had to be a change in the natural facts as well. I think Parfit would respond to this by saying that he is not denying that all of the natural facts determine all of the normatively *relevant* or *significant* facts, for example, what our psychologies are like, that we feel pain, that we have certain institutions, etc. What Parfit would deny is that all of these facts taken together capture or constitute normativity, in the reason giving sense. And if the world does not include normativity, then none of those facts really matter (Railton 2013 p. 1).

According to non-analytical, nonreductive naturalism, moral facts are *constituted* by natural facts, rather than being identical with natural facts. Parfit does not seem to consider this possibility. If moral facts are constituted or realized by natural facts this allows for the possibility of the multiple realizability of moral facts, meaning that although

in this world moral properties are “actually constituted or realized by natural properties” it did not necessarily have to be this way (Brink 1989 p. 158).⁸² On this view:

Moral facts and properties ...are constituted , composed, or realized by organized combinations of natural and social scientific natural facts and properties....This naturalist claim should be understood on the model of other common constitution claims: for instance, tables are constituted by certain combinations of microphysical particles, large scale social events and processes such as wars and elections are constituted by enormously complex combinations of smaller scale social events and processes, biological processes such as photosynthesis are composed of [physical events causally and temporally related in certain ways (Brink 1989 p. 159).

Another way of making this claim is to say that moral facts are a different level of description from natural facts, just as psychological facts are a different level of description than neuroscientific facts. Brink does not hang anything in his argument on the distinction between natural facts being identical to certain normative facts and natural facts constituting normative facts. But paying insufficient attention to this distinction may be what fuels some of Parfit’s intuitions about the incommensurability of normative and natural facts.

For example, in “The Normativity Objection,” Parfit argues that a natural “claim” and a normative “claim” could not possibly “state the same fact.” I assume here that Parfit is using “claim” to mean an expression or statement of a fact. I do not find this claim intuitively appealing, but Parfit seems to. Let’s let the natural fact be “this act maximizes happiness” and the normative fact be “this act ought to be done.” Parfit’s position sounds somewhat less compelling when put in terms of constitution. In that

⁸² Brink writes: Moral properties can be natural properties, though, even if they are not identical with natural properties. F can be G even if the property (or properties) designated by ‘F’ is not (or are not) the same as that (or those) designated by ‘G’. If G actually composes or realizes F, but F can be, or could have been, realized differently, then G constitutes, but is not identical with F...a table is constituted by, but not identical with, a particular arrangement of microphysical particles, since the table could survive certain changes in its particles or their arrangement (Brink 1989 p. 157-158).

case, the fact that the act maximizes happiness is what constitutes or realizes the fact that the act ought to be done.

3.4 Parfit's Triviality Objection to Non-Analytic Naturalism

Parfit's triviality objection against non-analytic naturalism relies on a distinction he makes between three types of normative claims, substantive, trivial, and positive. A normative claim is "substantive when these claims are significant, because we might disagree with them, or they might tell us something that we didn't already know" (Parfit 2011 p. 343). In contrast, trivial claims either do not contain any new information or are tautologies. Normative claims are positive "when they state or imply that, when something has certain natural properties, this thing has some other, different, normative property" (Parfit 2011 p. 343). Parfit ties the features of being a substantive and being a positive normative claim together in the following puzzling way: "When such claims are true, they state positive substantive normative facts" (Parfit 2011 p. 343). I will take Parfit to mean that for the class of normative claims that are true, in order for them to be claims that are philosophically interesting, they must also "state positive substantive normative facts" rather than being true but trivial, or true but whatever the opposite of a positive claim is.

In order to illustrate the distinction, he gives an example of what might seem like a moral view, but is in fact not. He imagines a utilitarian transplant surgeon who removes vital organs from one patient, killing him, in order to save the lives of a group of younger patients. When asked to justify his actions before the hospital's ethics committee he makes the following statement:

When I claimed that I ought to kill this patient, I was only stating the fact that this act would maximize happiness. On my view, that is the property to which the concept

ought refers. I was not claiming that this act would have some different property of being what I ought to do. On my view, there is no such different property. The property of maximizing happiness is the same as the property of being what we ought to do (Parfit 2011 p. 342).

Parfit claims that this statement of the surgeon's justification for his actions shows that the surgeon does not actually have a moral or normative view at all (Parfit 2011 p. 343).

Although Parfit is not explicit, presumably this is because the presumptively normative claim that the surgeon makes does not meet Parfit's conditions for a normative claim, the conditions of being both substantive and positive. The statement is not positive because the surgeon plainly denies that the natural "property of maximizing happiness" is anything different than being the "property of being what we ought to do" (Parfit 2011 p. 342). The condition for being positive states "when something has certain natural properties, this thing has some other, different, normative property" (Parfit 2011 p. 343). The surgeon's claim certainly seems substantive, in the sense that there will be people who do not share the surgeon's views.

Parfit uses two forms of utilitarianism to illustrate the triviality objection, a non-naturalist form of utilitarianism and a naturalist utilitarianism. He defines utilitarianism generically as the view that (A) "when some act would maximize happiness, this act is what we ought to do" (Parfit 2011 p. 341). According to Parfit (A) is a substantive moral claim, not a trivial claim. But (A) will be understood differently by non-naturalists and naturalists.

Non-naturalist utilitarianism says (B) "when some act would maximize happiness, this fact would make this act have the different property of being what we ought to do" (Parfit 2011 p. 341). Naturalist utilitarianism says (C) "when some act would maximize

happiness, that is the same as this act's being what we ought to do" (Parfit 2011 p. 341).

Parfit makes the following argument:

If, [...] (C) were true, (A) could not state such a fact. (A) could not be used to imply that, when some act would maximize happiness, this act would have the different property of being what we ought to do, since (C) claims that there is no such different property. Though (A) and (C) have different meanings, (A) would be only another way of stating the trivial fact that, when some act would maximize happiness, this act would maximize happiness.

Therefore: This form of Naturalism is not true (Parfit 2011 p. 343-344).

What is at issue here is whether or not the action has two different properties; the normative (ought) property and the distinct property of maximizing happiness. Or does the action have just one relevant property, the property of maximizing happiness which is the same property as the normative (ought) property. If (C) is the right way of understanding utilitarianism, then utilitarianism no longer makes a substantive moral claim, it makes merely a trivial one.

Finally, "[s]ince (A) is not trivial, (C) cannot be true" (Parfit 2011 p. 343-344). The reason Parfit thinks that any identity claim between a moral property and a natural property must be false, is that if it were true it would mean that what appear to be substantive moral claims (like A), would be trivial. Here he seems to be relying on the idea that property identity statements are always trivial because "every property is identical to itself" (Copp 2013). Frank Jackson helpfully recasts the argument in the following way:

Take any putative identification of being right with natural property N. If it is correct, saying that an act which is N is right will be trivial. But it manifestly isn't. For example, suppose the naturalistic candidate for being right is maximizing happiness, then [Parfit] argues, the claim that it is right to maximize happiness will be trivial. It can come to nothing more than the claim that 'being right is being right' or that 'maximizing happiness is maximizing happiness'. But no one thinks that. Many think that the claim that being right is maximizing happiness is false, and those who think it is true agree it needs argument, and, moreover, think it is

important to tell people about its truth. How then could the claim be trivial?
(Jackson 2012 p. 11).

A way of understanding the argument is as a dilemma that Parfit poses for the non-analytical naturalist, to suggest that their proposed identity between the normative facts and natural facts cannot be both positive and substantive at the same time (Railton 2013 p. 3-5). On the first horn of the dilemma, naturalism has to concede that there are actually two distinct properties, the normative property and the natural property, not one. This is because of the condition that Parfit establishes that the identity must be positive in order for it to count as a “true and informative” normative claim (Railton 2013 p. 4). Recall that “normative claims are positive when they state or imply that, when something has certain natural properties, this thing has some other, different normative property (Parfit 2011 p. 343). Thus any identity claim the naturalist makes will be self-defeating.

On the other horn of the dilemma, if naturalism gives a “referential reading of [the identity claim], so that it reaches directly to the underlying natural property, with no recourse to an intermediating normative property” this leads to the triviality of naturalistic identity claim (Railton 2013 p. 6).

3.5 Responding to Parfit’s Triviality Objection

Naturalists respond to this objection in at least three ways. The first has to do with the structure of the argument itself. The second, focuses on the nature of moral concepts and scientific concepts. The third response has to do with an assumption Parfit makes about the nature of identities. And the fourth response focuses on Parfit’s characterization of a positive normative claim.

Schroeder argues that as presented Parfit's argument for the triviality objection is not valid (Schroeder 2011 p. 6-7). This is because Parfit's condition for normative claim being positive says "when they state or imply that, when something has certain natural properties, this thing has some other, different, normative property" (Schroeder 2011 p. 6). The condition stipulates "some other, different, normative property" (Schroeder 2011 p. 7). This property does not have to be *same* normative property that is being reduced. Schroeder does not give an example, but I have tried to construct one here to clarify.

If the first normative property that we are trying to naturalize is "what we ought to do" and the natural property is "maximizes happiness," the only informative and positive normative claim need not be one that says "whatever maximizes happiness has another distinct normative property of being what we ought to do." We could instead say that: "whatever maximizes happiness is whatever we ought to do, and whatever maximizes happiness has *the other* normative property of maximizing rightness." This way the normative claim will have met the condition of being both substantive and positive. The property of maximizing right is another, different normative property, so it meets Parfit's condition of positivity.

A second way to respond to the triviality objection is to return to the property identifications that naturalist find analogous to the moral case, the identity between water and H₂O and heat and molecular motion. These identities are not trivial; in fact when scientists made these discoveries people acquired new information about the nature of water and heat.

However, Parfit anticipates this response when he argues that scientific concepts have gaps that are waiting to be filled, unlike normative concepts. Parfit's sense that

certain concepts like water and heat have gaps and that moral concepts do not, however, is just a sense he has. Parfit is aware that some kinds of concepts used to be thought to be in two “non-overlapping categories” (Parfit 2011 p. 324). But Parfit’s assertion that moral concepts are not in need of a natural explanation is a way of prejudging the outcome of our investigations. Methodologically, Railton argues, “it seems better...not to restrict the possibility of finding an underlying explanation or reduction of commonsensical categories by limiting this only to those concepts ‘ready-made’ for such explanation or reduction” (Railton 2013 p. 18). As Parfit is aware, there are historical examples of entities that at one time had “no preconceived ‘space’ for realization by purely physical systems” that we now believe science can give us reductive explanations of, like vital force (Railton 2013 p. 17).

Helpfully, Jackson’s distinction between a “role” and a “realizer” provides another convincing response to Parfit’s “gap” response. Jackson observes that most moral theories start with uncontroversial claims about the nature of right and wrong, good and bad, etc., what Jackson calls “folk morality” (Jackson 2012 p. 12).

The most attractive approaches to ethics start by making rather general claims that are urged to be self-evident, or *a priori*, or constitutive of the moral, or justified by the way they make sense of our most firmly held moral convictions, or....What’s right is what’s in accord with norms that no one could rationally object to...[etc.]...They [the general claims] give naturalists their riding instructions, as we might put it. The naturalists task, is then to find and identification of, let’s say being right, in terms of a natural property that makes true the general claim or claims in question...The implication of this picture is that any identification of being right will be an unobvious one. It takes real work to move from the general claim or claims to a thesis about which property fits the bill...(Jackson 2012 p. 12).

This response shows that Parfit’s claim that naturalist’s identifications of moral properties will not be trivial. This response is strengthened when we think about the

claims of folk morality that we begin our ethical theory with as giving us "... a job description that rightness...needs to satisfy" (Jackson 2012 p. 13). We do not know at the outset which property or set of properties will fulfill or satisfy that job description. Non-analytical naturalists think that it is possible to find the natural properties that satisfy the job description given to us by our starting assumptions ("folk morality"). Jackson calls the job description the "role property", and the kinds of properties that non-analytical naturalists hope to discover the "realizer property."

For example, if part of the job description gives a central place to what's desired in ideal circumstances, and allows that what's desired in ideal circumstances is contingent, then the property that fills the job description for rightness may vary from possible world to possible world. We will need in this case, distinguish the property that fits the bill-something that may vary from one world to another-from the property of being the property that fits the bill-something that is constant across worlds. The first is the realizer property, the second is the role property (Jackson 2012 p. 13).

There are many analogous cases in the sciences. Jackson uses the example of what it is to be a random sequence of numbers as being an important question in probability theory and statistics. Discoveries in that area provide us with new information, are controversial, and not trivial. Importantly, contra Parfit, a theory of what it is to be a random sequence does not mean that "there are *two* properties, that of being a random sequence and that of being so and so, where so and so is the account on offer-two distinct properties that are somehow glued together...There's one property" (Jackson 2012 p. 14).

Another way of responding to Parfit's insistence that naturalist identification of moral properties would be trivial, is to point out an assumption that Parfit makes about the nature of identities (Setiya 2011).

Parfit assumes, in effect, that when two concepts refer to a single property, and the first is unanalyzed, the content of the second can be explained in other terms. The challenge to the Naturalist is then to analyze normative concepts in a way that reveals the informational content of the alleged identity (Setiya 2011 p. 3).

So Parfit takes it for granted that if there is an identity between concept A and property X and there is also an identity between concept B and property X, this means that the two concepts, A and B, can be used interchangeably. But this is false. The morning star and the evening star are both Venus, but the *concepts* cannot be used interchangeably.

...the Naturalist ... should deny that the concept *should* be analyzed in terms of other concepts. It is informative to learn that what would satisfy one's final desires is what one should do, but not because one discovers that doing what satisfies one's final desires has a property distinct from that of being what one should do. One learns, precisely, that it is what one should do, information that cannot be expressed in any other way (Setiya 2011 p. 3).

If we are truly non-analytical naturalists, then we do not have to accept Parfit's challenge in the first place, that normative concepts can be analyzed in terms of naturalistic concepts.

More generally, we should question Parfit's definitions of substantive and positive moral claims (Railton 2013; Schroeder 2011). The condition that a normative claim must be positive in the manner that Parfit uses it is blatantly prejudices the outcome of any naturalist attempt at making identity statements between normative and natural facts. If the naturalist accepts Parfit's "positivity" condition, they have already given up the game. The terms of the debate, Parfit's conditions, do not seem to have an independent force unless one already rejects naturalism.

Parfit's arguments against non-analytical naturalism fail because he mischaracterizes the naturalists' position, the nature of the relationship between concepts and properties, and assumes a view on the un informativeness of the identities

that naturalists take to be out there to be discovered. For these reasons, Parfit's claim that naturalism cannot do justice to normativity does not succeed. Parfit's two major objections to non-analytic naturalism that were assessed here both failed to persuade that non-analytical naturalism is unequipped to handle the normative nature of moral facts. As others have noted, much of Parfit's discussion amounts to an insistence on his own certainty that a normative fact just cannot be the same thing as a natural fact.

4. Blackburn's Quasi-Realism

Simon Blackburn's quasi-realism is the final metaethical alternative to naturalist realism that I will consider. The main purpose in discussing quasi-realism is to use it as an illustration of the centrality of objectivity in explaining what a minimal account of the authority of morality could look like. I argue that because quasi-realism is not able to offer an adequate account of moral objectivity it certainly cannot capture the authority of morality. Following this, in the final section of the chapter, I argue that with respect to the concept of normativity or authority, many metaethical views are on equal footing with each other, but realism is able to capture the most important feature of morality's authority, its objectivity.

Because Parfit, Korsgaard, and others have argued that naturalist moral realism cannot capture the authority of morality in some sense, it is helpful to consider one more metaethical position that is, at least superficially, close to naturalist moral realism, and to consider whether that view can adequately capture the authority of morality.

Blackburn's rejection of naturalist moral realism functions in a different way from Parfit's or Korsgaard's. Blackburn is concerned with the excess metaphysical baggage

that any realist brings to the table. He argues that his leaner and meaner version of expressivism, quasi-realism, can capture the moral language and practice that realists claim to capture, without making any questionable metaphysical commitments.

4.1 Why Quasi-Realism?

Quasi-realism is Blackburn's attempt to make sense of certain aspects of moral phenomenology while avoiding the "queerness" of realism and without resorting to Mackie or Joyce style error-theory. According to Blackburn, the quasi-realist is "someone who 'starting from an anti-realist position finds himself progressively able to mimic the thoughts and practices supposedly definitive of realism" (Blackburn 1993 p.4). The moral phenomenology that Blackburn is referring to are the facts of our everyday moral practices, the way we speak about morality, and the role that morality plays in our life. In this context, we seem to assume that there is a "moral reality." For example, when we make moral judgments, although we tend to think we are making judgments that are correct, we admit the possibility that we could be mistaken in our judgment. Nor do we believe that thinking that something is morally wrong is what makes it morally wrong. Blackburn accepts that the moral phenomenology seems to assume some version of moral realism.

Blackburn thinks that the theory of quasi-realism, as a kind of projectivism posits the existence of fewer entities than moral realism, or than any version of cognitivism does, and that this is a theoretical advantage. Projectivism, is "the philosophy of evaluation which says that evaluative properties are projections of our own sentiments (emotions, reactions, attitudes, commendations)"(Blackburn 1984 p. 180). He claims

that quasi-realism does not have to posit a unique way of knowing or gaining access to moral facts, as a cognitivist may have to. Second, Blackburn argues that only projectivist quasi-realism can give a plausible account of the supervenience of moral properties on natural properties. Third, Blackburn argues that quasi-realism and projectivism offer a better account of moral motivation; if one accepts the Humean theory of motivation, which states that beliefs are inert and people only perform actions when beliefs are coupled with desires. If our moral judgments are beliefs, then the fact that they move us to act or refrain from acting in certain ways requires an appeal to a further desire to do the right thing or to be virtuous (See Chapters 2 and 3). If on the other hand, our moral judgments are expressions of non-cognitive states then they may have motivational force in themselves.

The moral appearances, then need to be reconciled with what Blackburn takes to be the only defensible account of the nature of moral judgments and moral thought. Moral judgments express non-cognitive emotional reactions to the world. This is a position about moral language, moral psychology, and the nature of moral properties. But at the same time quasi-realism seeks to explain why we can still say things like “it is true that stealing is immoral” and “whether something is right or wrong is not dependent on what we think about it.” In order to reconcile these two positions, Blackburn’s view relies on a deflationary account of truth and a projectivist metaethics.

4.2 Quasi-realist objectivity

The deflationary or minimalist account of truth that Blackburn relies on:

...allows us to end up saying ‘It is true that kindness is good.’ For this means no more than that kindness is good, an attitude that we may properly want to

express seems to let us end up saying, for instance, that 'kindness is good' represents the facts. For 'represents the facts' means no more than: 'is true'....we might even find ourselves saying that we know moral propositions to be true. Or really true, or really factually true, or really in accord with the eternal harmonies and verities that govern the universe, if we like that kind of talk (Blackburn 1998 p. 79).

We can consistently say that 'kicking dogs for fun is wrong' is true, or is really true, or represents the facts. We can speak this way without being committed to the existence of moral properties or facts in the metaphysically loaded way that the realist is. A deflationary or minimalist theory of truth is most easily understood in contrast to a correspondence theory of truth, which says that the propositional contents of our thoughts or sentences are true when they match the way the world is. The deflationary position on the other hands, says that:

...sentences ... express 'nondescriptive' propositional contents. These contents...do not purport to represent the world but rather play grammatical or logical roles very similar to those played by ordinary 'descriptive' propositions, such as embedding in propositional attitude descriptions (e.g., 'I believe that p') (Cuneo 2008 p. 179-180).

This means that the "nondescriptive propositional contents" of our sentences and beliefs are "neither are true nor correspond to correlative facts in [a particular] domain" (Cuneo 2008 p. 180-181). When applied to moral sentences and beliefs, the theory is supposed to leave the normal ways we speak about right and wrong untouched in their current forms, this is part of Blackburn's overall, quasi-realist strategy. He maintains that he is able to capture our everyday ways of thinking and speaking about morality just as well as the realist can. So on the deflationist view "...we can say that these contents are true or correspond to the facts...to say 'It is true that killing is wrong' is simply to repeat or endorse the (nondescriptive) proposition *that killing is wrong*" (Cuneo 2008 p. 180-181).

But if evaluative properties are “projections of our own sentiments (emotions, reactions, attitudes, commendations)” as Blackburn claims they are, then he has to be able to answer the person who says that if everyone felt differently than they currently do, the moral facts would be entirely different than they currently are (Blackburn 1984 p. 180). What can Blackburn say to the person who endorses the claim, ‘If we felt differently, then kicking dogs for fun would be morally right’ (Blackburn 1984 p. 218)? The person making the statement is expressing an attitude of approval for the attitude that ‘the wrongness of kicking dogs is dependent on our sentiments’ (Blackburn 1984 p. 218). That meta-attitude of approval is morally abhorrent, according to Blackburn.

For Blackburn, whether morality is objective or not is a first order normative issue, just as whether or not abortion is wrong, or whether people are endowed with universal human rights, are first order normative issues. This means that the question of whether kicking dogs for would be wrong regardless of whether anyone ever thought it was wrong, is a question from *within* morality rather than from *without*.

...on the construal of indirect contexts that I offer, it comes out as a perfectly sensible first-order commitment to the effect that it is not our enjoyments or approvals to which you should look in discovering whether bear-baiting is wrong (it is at least mainly the effect on the bear)’ (Blackburn 1993 p. 153).

Blackburn explains that it would be morally objectionable to hold the position that whether or not something is wrong depends on whether we think it is wrong. This is a moral view, according to Blackburn, about which features of the dog-kicking are the ones that make it wrong. The person who makes that statement is picking out our feelings or attitudes about dog kicking as the thing that makes dog kicking wrong, when what actually makes the dog kicking wrong is the pain it causes the dog (Blackburn 1984 p. 219). In other words, “...it is immoral to regard moral truths as mind-

dependent” (Rasmussen 1985 p. 187).

Blackburn takes his position to be justified because he understands the mind independence of morality as simply the denial of the “following counterfactual: had our sentiments been different, then the moral truth would have been different too”

(Rasmussen 1985 p. 187).

...what seems like a thought that embodies a particular second order metaphysic of morals is seen instead as a kind of thought that expresses first order attitude or need...on the construal of indirect contexts that I offer, it comes out as a perfectly sensible first-order commitment to the effect that it is not our enjoyments or approvals to which you should look in discovering whether bear-baiting is wrong...(Blackburn 1993 p. 153)

When faced with the question of what, other than the fact that I don't like it, makes kicking dogs wrong, “Blackburn ...den[ies] that there is any way to ask the question ... without asking ‘a moral question, with an answer in which no mention of our responses properly figures” (Brown 2008 p. 112). Blackburn relies here on Rudolph Carnap's distinction between internal and external questions, and claims that questions about the justification of our moral views are internal to our “moral framework” (Brown 2008 p. 112). In attempting to step into the domain of external questions to get a justification for a particular moral claim, according to Blackburn, “... we show that we are confused or some kind of anti-naturalist” (Brown 2008 p. 112-3).

The upshot is that it is consistent with quasi-realism that there to be “many moral frameworks” and if we try to determine which one of those frameworks is the correct (in the sense that “which framework corresponds to the way things are?”) one is asking an incoherent question (Brown 2008 p. 118).

Blackburn charges that if, as the realist would, one insists on getting an answer to the external question, about which one of the moral frameworks captures reality, this

shows that the realist must be relying on some version of a correspondence theory of truth that cannot be defended. Blackburn argues that he prevents a slide into relativism by pointing out that “there is no such thing a robust property of truth that would make [the external] question intelligible (Brown 2008 p. 118).

4.3 The problem with quasi-realist objectivity

When reading Blackburn’s quasi-realist explanation of the objectivity of morality it is easy to get the feeling that he is attempting to pull off some kind of a trick. This section will try to diagnose exactly what the trick is and why Blackburn’s account is not a satisfactory response to the person who claims that if everyone felt differently, kicking dogs for fun would be morally good (Rasmussen 1985; Zangwill 1994, 1990; Sinnott-Armstrong 2007).

A preliminary point is that Blackburn’s account amounts to a denial of one central aspects of metaethics. It denies that one of the central questions that have occupied metaethics is simply a mistake—that is the question of whether or not morality is fundamentally grounded in mind independent reality.

The first problem with Blackburn’s account is that there is a status difference between substantive moral claims and metaethical claims, which Blackburn ignores (Sinnott-Armstrong 2007; Zangwill 1994). Here I use the terms objectivity and mind independence interchangeably. I think that what is importantly at stake in the debate between quasi-realism and realism is objectivity, understood as the claim that moral facts remain the same regardless of what human beings think, believe, or feel about them. They are mind independent in this sense, not in the sense that if no minds were

to exist, moral facts would continue to exist unchanged, and not in the sense that moral facts have nothing to do with minds.

There is also a difference in status between the claims of morality at the level of moral theory or moral truths and the metaethical level, which have to do with conceptual and metaphysical truths (Sinnott-Armstrong 2007 p. 29). Claims about the rightness or wrongness of kicking dogs have a different status than claims about whether or not morality exists objectively in a mind independent way. Zangwill argues that analogously, if someone held the position that whether or not $2+2=4$ is dependent on our attitudes about their sum, this would not count as a “mathematical attitude or belief,” akin to the belief that $1 \times 3=3$. This would have to be stated at some other level of belief about the nature of mathematics. And Zangwill wonders, if an “internal reading” of the meta-mathematical claim” does not seem correct then neither should Blackburn’s internal reading of claims about moral objectivity and mind independence make sense either (Zangwill 1994 p. 209-210).

Zangwill also offers an account of what is significantly different about the way realists and quasi-realists like Blackburn talk about objectivity and mind independence. One of the suggestions for why there is a difference in conceptual status between these claims is the claim that mind independence has “generativity,” which means that:

...the principle of mind-independence can be instantiated by infinitely many particular cases. Mind-independence is the principle that 'It is not the case that the wrongness of anything depends on my attitudes'. Quasi-realism needs to achieve such a general formulation and not merely a formulation in terms of kicking dogs (Zangwill 1994 p. 210).

Part of Zangwill’s point here is that from a metaethical concept like mind independence, we can infer that any case of wrongdoing is independent of our attitudes (Zangwill 1994

p. 210). Quasi-realism's mind independence has to be articulated on a case by case basis.

The second problem with Blackburn's account is that it does not really capture our moral practice and everyday ways of moral thinking about objectivity in the same way that realism can (Sinnott-Armstrong 2007 p. 29). This is a particular problem for Blackburn because one of the virtues of quasi-realism is supposed to be that it can capture all of the same practices that realism can. When we moralize, when we make moral judgments, we are aiming at or aspiring to making moral judgments that are correct. This also happens when we disagree about morality with other people. The possibility of a moral judgment being incorrect is usually thought to have the status of a conceptual truth. A system of moral judgments without the assumption of correctness built into them would just be a set of judgments about the niceness of people, acts, and desires, but not a system of moral judgments.

...Blackburn glosses the counterfactuals as claims about which patterns of sensibilities are or are not acceptable, but people take themselves to be talking about kicking dogs rather than about attitudes when they assert such claims or counterfactuals. If some people fail to disapprove of kicking dogs but still do not perform acts of kicking dogs, then I might think less of them, but at least they do not do any of the acts that I am talking about that I call 'when' when I say, 'Kicking dogs is wrong.' The same point applies to the counterfactual, "Even if we all approved of it, kicking dogs would still be wrong.' The consequent is a judgment about acts in another possible world, not about attitudes in this world or that. Thus even if Blackburn does capture one kind of mind independence, it is different from the kind of mind-independence that realists find implicit in common moral thought and talk (Sinnott-Armstrong 2007 p. 29).

What is important for preserving what a realist thinks that we want from moral mind independence is modal independence of some kind (Jenkins 2005). When we say that right and wrong are independent of what we think about them, we mean that certain acts remain right or wrong in other possible worlds, in which people have different

beliefs about them. Whereas Blackburn's version of mind independence is really about the attitudes of other people, people who say things like, 'right and wrong just depend on what we think about them'.

There are good reasons to doubt that quasi-realism captures moral objectivity, in the sense of mind independence, despite Blackburn's reply that to think that moral judgments are relative to what a particular person or society thinks about them is a matter of first order ethics.⁸³

Blackburn's quasi-realism does not seem to be able to give a reason to think that any individual or group's opinions, beliefs, or emotive reactions to states of affairs are justified or even more justified than any others. Although Blackburn (and most others) thinks that torturing dogs for fun is wrong, when pressed to give a justification for this reaction, or one of the more fundamental reactions underlying it, he cannot. What Blackburn fails to provide, Brown argues is, "an account...of why it is that we *ought* to feel one way rather than another" about a particular moral issue (Brown 2008 p. 104).

The projectivist interpretation, therefore, is another attempt to represent the rational criticism of attitudes by appealing to higher-order attitudes. But the multiplication of attitudes does not capture the possibility of the criticism or evaluation of attitudes; it always leaves some attitude beyond criticism that apparently ought not to be beyond criticism (Irwin 2009 p. 834).

The problem here is essentially that Blackburn will either face an infinite regress of attitudes or he will have to say that at some point, certain moral attitudes we have are foundational, that is they are not justified in terms of another "higher-order attitude." If

⁸³ It would be too far afield to go into a detailed discussion of why Blackburn's reliance on a deflationary account of truth fails to secure him the justification he needs. Briefly, Blackburn assumes not only a "redundancy theory of truth" but also a "deflationary theory of truth". The first is a theory about the meaning of "is true", that is, "to say that p is true is to say no more than p"; the deflationary theory of truth "says that there is nothing more to the property of being true than the T-schema" (Brown 2008 p. 105). And the deflationary theory of truth "cannot save Blackburn from the charge that quasi-realism collapses into mere autobiography" (Brown 2008 p. 106).

this is the case those attitudes that are foundational are either unjustified or arbitrary. The alternative is, as a realist would contend, that they are justified in terms of their relationship with some kind of moral reality that is independent of those attitudes. Blackburn's interpretation of mind independence means that it is only possible to "...disapprove of some attitudes by reference to other attitudes" but, "this is not how we normally understand the fallibility of attitudes" (Irwin 2009 p. 834).

5. Moral Realism and a Minimal Understanding of the Authority of Morality

The objection to naturalist moral realism based on the authority of morality says the authority that morality has over us is incompatible with some element of naturalist moral realism, either the element of the theory that makes it realism, or the element of the theory that makes it naturalist.

I have responded to this objection in several steps. First I have considered some of the most frequently discussed accounts of what the authority of morality could consist in inescapability, overridingness, categoricity, bindingness, or normativity. A recurring theme of this work is to distinguish these accounts of the authority of morality from the motivational nature of our moral judgments. The way that our moral judgments motivate us are psychological facts about human beings. No matter how motivational our moral judgments turn out to be, it is always appropriate to ask whether or not we should be motivated to act on them. This is the case with any kind of consideration that may provide us with motivation. A lot of confusion regarding the nature of morality's authority can be avoided once the difference between authority and motivation is appreciated. One of the positive preliminary claims that naturalist moral realism can make about the

authority of morality is that authority is a feature of morality itself, rather than a feature of human beings. Authority does not have to do with our motivational system, whether that is construed in terms of reasons or desires.

After the motivational issue has been tabled, it becomes more difficult to see why naturalist realism would have a harder time than other metaethical theories accounting for any of the various plausible interpretations of the authority of morality. What kind of authority does morality then have on the naturalist moral realists view? If you ignore the demands of morality, you are probably morally bad. Morality applies to you whether you care about it or not, in the same way gravity applies to you whether you care about it or not. It seems like this is a striking dissimilarity. You cannot bump into moral properties in the way that you can bump into walls. Of course, if you ignore the laws of gravity you might fall down and get hurt. If you ignore the laws of morality, the moral facts and properties, you will not fall down. But you will be a bad person, and other people will probably notice. The world and your identity change when you ignore moral facts and properties, the properties that we can ascribe to you also change. Of course, what makes a property or fact respectably real is not the fact that if you ignore it something bad will happen to you.

Naturalist moral realism by itself will probably not be able to provide an answer to the normative moral skeptic who demands to know why they should do what they are morally required to do and is not satisfied by the reasons provided by the appropriate first order normative theory. But no other metaethical theory can secure this answer either. What naturalist moral realism can provide that brings it closer to accounting for the authority of morality that other metaethical positions cannot is the objectivity or mind

independence of morality. Presumably, for a set of considerations to be authoritative, it is at least required that they be true, real, or objective (FitzPatrick 2005). This is the minimal account of the authority of morality, authority as objectivity.

A prominent realist strategy for understanding the authority of morality is shifting this discussion to one about practical reason or rationality. Realists then argue that in order to reason well, one has to recognize moral reasons as paramount. Or they may argue that moral considerations necessarily give us reasons for acting, even if they do not necessarily motivate us to act on them. I have chosen not to directly engage in a debate about reasons or rationality in this chapter because I do not think that talk of reasons can secure the authority of morality for realism. In part, this is because talking about reasons has the drawback of either being understood as another purely normative discourse or a folk psychological one.

On the one hand, many acknowledge that theories of practical reason and rationality are, like theories of morality, also a normative enterprise. If the issue of where the authority of morality comes from and what the authority of morality is gets translated into talk about reasons, then the question of the authority is merely being postponed, passed off to another normative domain. The issue will inevitably resurface in the form of a question about the authority of practical reason. Alternatively, rationality and practical reason can be understood as purely descriptive theories of either folk psychology or to be investigated by psychology and cognitive science proper. But if that is the case, it is unlikely that those theories will be of any help in securing the authority of morality. Psychological theories, whether folk or scientific, can tell us about ourselves; they can give us insight into the truth of cognitivism or non-cognitivism, and

the nature of moral motivation. But it is not clear how these theories will help to tell us anything new about the authority of morality itself.

Korsgaard and Parfit write about the authority of morality in terms of the normativity of morality. Korsgaard argues that it is realism that cannot secure the normative force of morality essentially because it is a metaphysical rather than a practical project. Many of her objections to realism fail on their own grounds. But more importantly, she sets a standard for being able to answer the normative question that her theory fails to meet. At the same time, Korsgaard's Kantian constructivism is unable to secure an objective status for morality.

Korsgaard forcefully posed the normative force objection by saying that a metaethical theory that deals with the objection must be able to give an answer to the skeptical question 'why should I do what I am morally required to do?' Although Korsgaard's own theory does not provide an adequate answer to this question, it might still be that an answer is in fact a requirement of any metaethical theory. Parfit's theory of normativity cannot answer this question either. His answer is that we must do what we have reason to do. But as with Korsgaard's theory, there is always the possibility for a regress here, why should I do what I have most reason to do?

Parfit uses normativity in a different way than Korsgaard does and focuses on naturalism's inability to explain ultimate normative reasons. However, Parfit's arguments against non-analytic naturalism fail because they either misunderstand what a non-analytic naturalist is really committed to or they simply beg the question, asserting that no identity could ever be discovered between natural facts and moral facts that could

preserve the reason giving force of the moral facts. Parfit's theory is further complicated by the fact that moral reasons themselves are *sui generis* and unanalyzable.

The last theory considered was Blackburn's quasi-realism. Some may have been convinced with the responses given to Korsgaard and Parfit and yet still suspicious of realism's ability to deal with the authority or normativity of morality. If these suspicions stem from a general suspicion of the ontology of realism, they may be sympathetic to a theory like Blackburn's because it seems to get all of the authority, objectivity, and truth we are looking for in a metaethical theory, without any unsightly metaphysical baggage. The discussion of quasi-realism in this chapter was by no means a full analysis of the assets and liabilities of that theory. Instead, it focused on quasi-realism's ability to secure moral objectivity and justification because they are preconditions for the ability to secure the authority or normativity of morality. As we saw quasi-realism falls short on these grounds.

Quasi-realism cannot adequately answer the skeptical question. And as Zangwill and Brown argue, Blackburn cannot justify one set of emotional reactions to an action or state of affairs over another. Although Blackburn does not address moral skepticism directly, we can imagine what he might say to the skeptic based on his response to the relativist who claims that whether or not kicking dogs is wrong or not depends on what an individual or some group of people thinks about it. So when asked by the skeptic why I should do what I am morally required to do? Blackburn can only tell him that the question only makes sense from within a moral framework, and that if he does not think he should do what he is morally required to do he is a morally deplorable person.

Conclusion

Moral realism is the view that moral judgments either succeed or fail in corresponding to a mind independent moral reality. The moral facts about people, actions, states of affairs, and so on and the moral properties of these people, actions, states of affairs, exist regardless of what we think of them or whether or not we are even aware of them. Moral realism has strong common sense appeal, so much so that it can be assumed to be the default metaethical position. This project has been primarily a defense, rather than a positive account of moral realism.

In contemporary metaethics one of the central worries about moral realism is that unless realism can be reconciled with naturalism, realism is left open to the charge that talk of objective moral facts and properties should go the way of talk about phlogiston and witches, which seemed to be common sense at one point, but have since been abandoned as metaphysically and epistemologically indefensible.

A commitment to naturalism, whether methodological, epistemological, or metaphysical, may seem incompatible with moral realism because moral realism might be thought to embrace a realm of *sui generis* or mysterious moral facts and properties and a corresponding faculty with which to perceive them.

The position I have defended is a version of Cornell realism. Cornell realism is the view that moral properties and facts exist objectively; this means that they are not dependent on or constituted by what we think about them. At the same time, these moral properties and facts are constituted by basic physical properties and facts (See Introduction and Chapter 1, Section 1). Cornell realism directly addresses some of the perpetual worries about the incompatibility of realism and naturalism by characterizing

moral facts and properties as being identical with, constituted by, or supervening on basic physical facts and properties, in an analogous way to the relationship between the psychological or historical facts and properties and basic physical facts and properties.

Another virtue of Cornell realism is that it posits non-analytic identities between moral properties and natural properties. Rather than having to do with truths of meaning, these identities are to be understood along the same lines as the scientific property identities that we have discovered between things like water and H₂O.

The version of Cornell realism I have defended is neutral on whether or not the identities between moral properties and facts and natural properties and facts are reductive or nonreductive. Nonreductive realism says that that moral properties and facts are constituted by natural properties and facts, rather than being reduced to natural properties and facts. While I favor nonreductive realism, that debate has been largely avoided in this essay.

Despite the plausibility of Cornell realism, queerness objections, articulated influentially by J.L. Mackie to realism itself and naturalistic realism, specifically the existence of moral facts and properties have remained. Mackie's set of queerness objections to the existence of objective values continue to be cited and contemporary challenges to realism grounded in the queerness or moral facts and properties abound. The two versions of queerness objections that I have spent the most time addressing are objections based on moral motivation (the internalism objection) and objections based on the authority or normativity of morality.

I have argued that Cornell Realism can withstand these two central objections. A theme throughout this work has been pointing out arguments and positions that depend

for their plausibility on either subtly or blatantly conflating questions of how morality motivates with questions about the nature of the authority that morality has over us. When these two objections are properly distinguished from each other and other sorts of queerness objections, moral realism is well positioned to account for the nature of moral motivation, especially if it embraces a pluralistic externalism about motivation.

In Chapter 1, I introduced Mackie's classic queerness objections to the existence of objective values and I cataloged his several sub-arguments (Chapter 1, Section 1). Mackie argues that if objective values were to exist, they would be a kind of thing like nothing else that exists. Cornell realism succeeds in addressing one element of Mackie's worry. It addresses the assumption that moral properties and facts (objective values, in Mackie's language) cannot be reconciled with a world of natural properties and facts. However, even if Cornell realism succeeds in this capacity, I argue that the view must still contend with a set of challenges derived from Mackie's queerness objections.

After presenting some of the central ways that the queerness objection can be understood, I focused on the strain of the argument having to do with metaphysical queerness (Chapter 1, Section 2). Metaphysical queerness itself can be understood in several different ways. The metaphysical queerness of moral properties and facts might have to do with explaining the contentious relationship of supervenience which is said to hold between moral properties and facts and natural properties and facts. Or it may have to do with the concern that objective values (moral properties and facts) do not seem to be causally efficacious in the way other properties and facts may be. Finally, it can be understood to have to do with the normative force of objective values. In the

remainder of the essay I focused on the normative force strain of Mackie's argument. Mackie's objections that center on normative force contain at least two sub objections. I further parse the objection into an objection having to do with the nature of moral motivation and an objection having to do with the authority of morality.

I argued that these objections are often conflated. But, in fact objections to realism based on concerns about the nature of moral motivation are distinct from objections to realism based on the worry that it cannot capture the authority of morality. Both objections claim that there is something about realism (natural or non-natural) that makes it unable to easily accommodate or be consistent with an important belief that we have about the way morality works or an important fact about morality. A central aim of this essay has been to distinguish these two objections from each other and to provide a plausible way of understanding each objection as a problem that realism must face.

The objection based on the nature of moral motivation says that because realism is committed to cognitivism it has to construe moral judgments, either the sentences we utter with moral content or the mental states we have with moral content, primarily as beliefs. The objection claims that beliefs, however, are motivationally inert; they cannot motivate us to act on them all by themselves. But moral judgments do motivate us to act all by themselves, the objection continues. So cognitivism is incompatible with a significant piece of moral psychology and phenomenology. The view that the objection relies on here is motivational judgment internalism (internalism for short). Realism is therefore probably not true because the cognitivism it is committed to is incompatible with internalism. A more plausible metaethical view, the argument continues, can allow for motivation to be built into moral judgments themselves, candidates include

expressivism, projectivism, emotivism, Smith's unique rationalism, and so on (Chapter 1, Section 3).

The second queerness objection that I contend with in the essay has to do with the authority of morality. It says that realism (whether naturalistic or not) cannot capture, or is incompatible with, the unique way that morality has authority over us. What this kind of authority comes down to is something that I explore in Chapter 4. Initially it seems that morality is not the sort of thing one can opt out of; it binds us all regardless of our wishes to the contrary. On a realist view, the objection claims, the "ought-ness" of morality, even morality's normativity disappears or becomes incoherent. Korsgaard's formulation of this objection identifies realism itself as the problem, whereas Parfit's formulation of the objection identifies naturalism as the source of the problem (Chapter 2, Section 4).

Chapter 1 included a review of and responses to several of the most historically significant objections to moral naturalism which have a continued influence on contemporary metaethics. I presented the problems for naturalism posed by David Hume's alleged is-ought gap and G.E. Moore's naturalistic fallacy. I made several arguments against these classic objections to moral naturalism (Chapter 1, Sects. 5 and 6).

Chapter 1 concluded with a brief discussion of a final argument that Mackie makes against objective values, the objection from disagreement. Mackie distinguishes from the queerness objection, but the objection is often raised against realism in conjunction with the queerness objections, so it merits addressing (Chapter 1, Section 7).

In Chapter 2 I addressed the internalism objection to realism; this is the objection that has to do with moral motivation. Internalism is the view that moral judgments necessarily, inherently, or essentially motivate the agents who make them to act on them. The competing view is externalism, which says that moral judgments do not necessarily, inherently, or essentially motivate the agents who make them to act on them. According to externalism, there is a contingent connection between making a moral judgment and having a moral motivation. I began by giving a positive case for internalism and outlining several different types of internalism (Chapter 2, Sections 1 and 2).

I considered the positive arguments in favor of internalism, the argument from phenomenology, the argument from oddness, and the argument from connection. I argued that all of these arguments face serious challenges. I gave particular attention to the connection argument and a particularly plausible version of it put forward by Smith. In addition to responding to the arguments in favor of internalism I argued that there are independent reasons to doubt internalism.

Finally, I developed a plausible alternative to internalism which I call a pluralistic externalist account of moral motivation. At times, we are motivated to act on our moral judgments by sources like compassion and sympathy. We are also motivated to act on our moral judgments by conative states like desires, which have content related to the moral judgment we make. For example, I might desire to do the right thing. Another person or at another time, someone may desire that their friend be happy. Another time a person may act out of a mostly or purely non-cognitive affective state like sympathy for the suffering person (Chapter 2, Section 3).

I argue that pluralistic externalism satisfies three important conditions that any theory of moral motivation should meet, the conditions of flexibility condition, the neutrality condition, and consistency with empirical evidence.

One of the most attractive features of pluralistic externalism is that it can easily accommodate the conceivability and existence of amoralists and other forms of moral indifference. Pluralistic externalism is a theory of moral motivation that is flexible enough to be compatible with the wide range of variability that we find in moral motivation between individuals and within a single individual across times and situations. Pluralistic externalism can embrace a wide range of types of motivation and can thus also easily explain a wide range of ways in which one can fail to be motivated as well.

A theory of moral motivation should also be impartial with respect to ethical theories about what kind of motivation is morally praiseworthy; it should be as normatively neutral as possible. Thus the third condition is the neutrality condition. Pluralistic externalism is agnostic on which type of motivation represents the most morally, or the only morally praiseworthy type of motivation. Further, it is neutral about whether or not motivation is something that can be or should be morally assessed at all or whether it is rather the action that comes out of our deliberative process that truly matters morally. Neutrality on this issue is important even if one does not accept a strict division between metaethics and normative ethics. I do not embrace such a strict division but I take the truth and nature of a theory of moral motivation to be an empirical theory of moral psychology. The way that our motivational system actually works is likely relevant to what should be counted as morally praiseworthy or worthwhile. But this

does not mean that moral psychology should build normative views into the theories of moral motivation themselves.

Finally, a theory of moral motivation should be consistent with the existing psychological and neuroscientific evidence about how the human motivational system works in general (the empirical evidence condition). Although in this essay I have not presented empirical evidence in any detail, I argue that because pluralistic externalism is an *a posteriori* theory, as opposed to most versions of internalism, which are *a priori*, pluralistic externalism is open to being amended based on scientific evidence.

In Chapter 3 I delved in the phenomenon of moral indifference in detail. I presented some of the most important and widely discussed forms of moral indifference and amoralism. The morally indifferent, like amoralists, are people who appear to be able to make moral judgments without being motivated to act on those moral judgments. The mere conceivability of such people is a serious counterexample to *a priori* internalism. Their conceivability does not affect the less common *a posteriori* forms of internalism. If the morally indifferent exist and if the externalist characterization of them can be defended, then the counter example touches *a posteriori* forms of internalism as well (Chapter 3, Section 1).

I argued that moral indifference is common. This becomes clear when we stop considering only the monstrous figure of the amoralist and broaden our scope to consider the variability in levels of moral motivation between individuals and across an individual over time. I addressed the ways that internalists attempt to respond to the conceivability and existence of moral indifference. I especially considered the internalist strategy of reinterpreting what the morally indifferent person says to show that they do

not really make a moral judgment at all. The internalist responses to moral indifference were all found lacking and concluded that the phenomenon of moral indifference is something that the best theory of moral motivation must be able to account for and argued that pluralistic externalism can do just that (Chapter 3, Sects. 2 and 3).

The authority objection to realism was addressed in Chapter 4. When we try to understand what the authority objection is really about we find several different strains of thought, inescapability, overridingness, bindingness, and normativity, none of which is incompatible with a realist picture of the nature of moral facts and properties. A realist account of morality's authority is able to uniquely capture the objective nature of morality in a way that many other metaethical theories cannot. Since naturalism can at the same time resist positing a *sui generis* realm of moral properties it is a more convincing metaethical position (Chapter 4, Section 1).

The authority objection to naturalist moral realism is motivated by several distinct metaethical positions. One way of understanding the objection is that if we claim that moral properties are mind independent features of the world, we lose the central connection that morality has to guiding our decision making. Although this authority objection comes close to the motivation objection, the objection is not meant to be understood in a psychological or motivational sense. Since there are several different ways the authority of morality has been interpreted, I considered two specific objections to naturalist moral realism based on the claim that it cannot capture the authority of morality.

I first considered an objection from Korsgaard, who focuses on realism as a metaphysical theory. She argues that realism does not have the resources to answer

what she calls the normative question; it cannot explain why an objective realm of mind independent facts should have any implications for how we should reason or behave. I offered independent reasons against Korsgaard's objection to realism. I also argue that her theory, Kantian constructivism, cannot meet the standard that requires that realism meet (Chapter 4, Section 2).

Chapter 4 explained that, for Korsgaard, every metaethical theory must be able to explain, in a first person context, why someone should do what is morally required of them. Korsgaard claims that while realism fails to sufficiently answer the normative question, her metaethical position, Kantian constructivism succeeds. I responded to Korsgaard's objection to realism. First, Korsgaard does not consider or unfairly dismisses the kinds of answers that a realist can give to the normative question. Second, Korsgaard's argument from the normative question sometimes conflates justification and motivation; that is, it combines the internalism objection and the authority objection in subtle ways. I also argued that Korsgaard's normative question challenge to moral realism is equally problematic for her own Kantian constructivism. More broadly, I claimed that we have reason to suppose that the way she frames the challenge all metaethical theories will face difficulties answering her normative question. However, realism, however, fares better than many others on the grounds that it captures objectivity in a unique way.

Second, I discussed Parfit's version of the authority objection to naturalist versions of realism, specifically, non-analytical naturalism. I presented several responses to Parfit's objection to naturalism. Parfit's objection to all naturalistic metaethical theories, including what he calls non-analytical naturalism, is that

naturalistic views eliminate the reason giving sense of normativity. The only metaethical view that can capture the reason giving sense of normativity, Parfit argues is a nonnatural realism that includes a *sui generis* realm of nonnatural normative facts. Parfit provides three arguments against non-analytical naturalism, “the normativity objection,” “the triviality objection,” and “the fact-stating argument.” The goal these three arguments share is that they attempt to show how that natural facts by themselves cannot be normative. Parfit’s arguments against non-analytical naturalism fail for several reasons. Simply, Parfit’s arguments beg the question against non-analytical naturalism. They persuade only if one begins by assuming that normativity has to be non-natural (Chapter 4, Section 3).

Finally I discussed Blackburn’s metaethical theory quasi-realism, which he presents as a view that can capture the desirable features of moral realism, such as the objectivity of morality and the phenomenology of moral disagreement and reasoning without expanding his ontology to include objective moral properties or facts. I do not give a full account of Blackburn’s view. Instead I sketched the view as a way of highlighting the importance of the objectivity of morality for a minimal account of the authority of morality by contrasting the kind of objectivity that moral realism can secure, with the kind of pseudo-objectivity that non-realist theories, like Blackburn’s can secure (Chapter 4, Section 4).

Having argued that neither Korsgaard’s or Parfit’s paradigmatic cases of the authority objection to naturalist moral realism succeed, I closed Chapter 4 with a discussion of the kind of authority that naturalist moral realism does have. The authority of morality is not the same as objectivity. But in reviewing Blackburn’s position, it

becomes easier to see the kind of authority that realism can capture that non-realist views, like constructivism or quasi-realist expressivism cannot. I end the chapter by discussing what kind of authority is the best we can hope for, that is objectivity. This is something the realist can secure. This is a particular success for naturalist moral realism if at the same time it is able to do all of this without positing the existence of *sui generis* moral properties or facts (Chapter 4, Section 5).

References

- Alter, Torin. 1998. A limited defense of the knowledge argument. *Philosophical Studies* 90 (1):35-56.
- Archard, David. 2011. Why moral philosophers are not and should not be moral experts. *Bioethics* 25 (3):119-127.
- Arcuri, Luciano, Luigi Castelli, Silvia Galdi, Cristina Zogmaister, and Alessandro Amadori. 2008. Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology* 29 (3):369-387.
- Ayer, A.J. 1936/2012. *Language, Truth, and Logic*. New York: Dover Publications.
- Bagaric, Mirko. 2002. Internalism and the part-time moralist: An essay about the objectivity of moral judgments. *Consciousness & Emotion* 2 (2):255-271.
- Bambrough, Renford. 1969. Proof of the Objectivity of Morals, *American Journal of Jurisprudence*. 14:37.
- Baumrin, Bernard H. 1968. Is There a Naturalistic Fallacy? *American Philosophical Quarterly* 5 (2):79-89.
- . 1986. Moral Blindness. *Metaphilosophy* 17 (4):205-213.
- Bedke, Matthew S. 2009. Moral judgment purposivism: saving internalism from amoralism. *Philosophical Studies* 144 (2):189-209.
- . 2010. Might All Normativity Be Queer? *Australasian Journal of Philosophy* 88 (1):41-58.
- . 2012. Against Normative Naturalism. *Australasian Journal of Philosophy* 90 (1):111-129.
- Belmore, M.F., and V.L. Quinsey. 1994. Correlates of psychopathy in a noninstitutional sample. *Journal of Interpersonal Violence* 9 (3):339-349.
- Björnsson, Gunnar. 2002. How emotivism survives immoralists, irrationality, and depression. *The Southern Journal of Philosophy* 40 (3):327-344.
- . 1998. *Moral Internalism: An Essay in Moral Psychology, Philosophy*, Stockholm University.
- . 2012. Do 'Objectivist' Features of Moral Discourse and Thinking Support Moral Objectivism? *The Journal of Ethics* 16 (4):367-393.
- Björklund, F., G. Björnsson, J. Eriksson, R.F. Olinder, and C. Strandberg. 2012. Recent Work on Motivational Internalism. *Analysis* 72 (1):124-137.

- Björnsson, Gunnar, and Ragnar Francén Olinder. 2013. Internalists Beware-We Might all be Amoralists! *Australasian Journal of Philosophy* 91 (1):1-14.
- Blackburn, Simon. 1971. Moral Realism. In *Morality and Moral Reasoning*, edited by J. Casey. New York: Routledge.
- . 1984. *Spreading the Word*. Oxford: Clarendon Press
- . 1988. Supervenience Revisited. In *Essays on Moral Realism*, edited by G. Sayre-McCord. Cornell, New York: Cornell University Press.
- . 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- . 1998. *Ruling Passions*. Oxford: Clarendon Press
- Blair, R.J.R. 1995. A cognitive developmental approach to morality: Investigating the psychopath. *Cognition* 57 (1):1-29.
- Bostrom, Nick, and Toby Ord. 2006. The Reversal Test: Eliminating Status Quo Bias in Applied Ethics. *Ethics* 116 (4):656-679.
- Boyd, Richard N. 1988. How to be a Moral Realist. In *Contemporary Materialism*, edited by P. K. Moser and J. D. Trout. London: Routledge.
- Brink, David O. 1997. Kantian rationalism: Inescapability, authority, and supremacy. *Ethics and Practical Reason*. Edited by G. Cullity and B.Gaut. New York: Oxford University Press:255-291.
- . 1984. Moral realism and the sceptical arguments from disagreement and queerness. *Australasian Journal of Philosophy* 62 (2):111-125.
- . 1986. Externalist moral realism. *The Southern Journal of Philosophy* 24 (S1):23-41.
- . 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge Univ Press.
- . 1997. Moral motivation. *Ethics* 108 (1):4-32.
- Brown, Charlotte. 1988. Is Hume an internalist? *Journal of the History of Philosophy* 26 (1):69-87.
- Brown, Richard. 2008. The Semantics of Moral Communication. Dissertation, Philosophy, The Graduate Center, City University of New York, New York.
- Butler, J. 1726. *Fifteen Sermons Preached at the Rolls Chapel*. London: W. Botham.
- Cholbi, Michael. 2006. Belief attribution and the falsification of motive internalism. *Philosophical Psychology* 19 (5):607-616.

- . 2011. Depression, listlessness, and moral motivation. *Ratio* 24 (1):28-45.
- Cima, M., F. Tonnaer, and M.D. Hauser. 2010. Psychopaths know right from wrong but don't care. *Social Cognitive and Affective Neuroscience* 5 (1):59-67.
- Cohen, G.A. 1996. Reason, humanity, and the moral law. *The Sources of Normativity*. Cambridge: Cambridge University Press:167-188.
- Copp, David. 1995. Moral obligation and moral motivation. *Canadian Journal of Philosophy* 25 (sup 1):187-219.
- . 1995. *Morality, Normativity, and Society*. New York: Oxford University Press.
- . 1997. Belief, Reason, and Motivation: Michael Smith's "The Moral Problem". *Ethics* 108 (1):33-54.
- . 2007. *Morality in a natural world : selected essays in metaethics, Cambridge studies in philosophy*. Cambridge; New York: Cambridge University Press.
- . 2013. Normativity and Reasons: five arguments from Parfit against normative naturalism. In *Ethical Naturalism: Current Debates*, edited by S. Nuccetelli and G. Seay. New York: Cambridge University Press.
- Cuneo, Terence D. 1999. An Externalist Solution to the "Moral Problem". *Philosophy and Phenomenological Research*:359-380.
- . 2008. Moral Realism, Quasi Realism, and Skepticism. In *The Oxford Handbook of Skepticism*, edited by J. Greco. New York: Oxford University Press.
- Damasio, H., T. Grabowski, R. Frank, A.M. Galaburda, and A.R. Damasio. 1994. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264 (5162):1102-1105.
- Dancy, Jonathan. 1993. *Moral Reasons*. Oxford: Blackwell.
- . 2006. Nonnaturalism. In *The Oxford Handbook of Ethical Theory*, edited by D. Copp. New York: Oxford University Press.
- Darwall, Stephen. 1992. Internalism and agency. *Philosophical Perspectives* 6:155-174.
- . 1997. Reasons, motives, and the demands of morality: An introduction. In *Moral discourse and practice: Some Philosophical Approaches*, edited by S. Darwall, A. Gibbard and P. Railton. New York: Oxford University Press.
- Darwall, Stephen, Allan Gibbard, and Peter Railton. 1992. Toward fin de siecle ethics: some trends. *The Philosophical Review* 101 (1):115-189.
- Devitt, Michael. 1996. *Coming to Our Senses: A Naturalistic Program for Semantic Localism*. New York: Cambridge University Press.

- . 2002. Moral realism: A naturalistic perspective. *Croatian Journal of Philosophy* (4):1-15.
- . 2010. *Putting metaphysics first: Essays on metaphysics and epistemology*: Oxford: Oxford University Press.
- Dreier, James. 2000. Dispositions and Fetishes. *Philosophy and Phenomenological Research* 61 (3):619-638.
- Dummett, Michael. 1996. *Origins of Analytical Philosophy*. Boston: Harvard University Press.
- Enoch, David. 2007. An outline of an argument for robust metanormative realism. In *Oxford Studies in Metaethics*, edited by R. Shafer-Landau. Oxford: Oxford University Press: 21-50.
- . 2009. How is moral disagreement a problem for realism? *The Journal of Ethics* 13 (1):15-50.
- Falk, W David. 1947. "Ought" and Motivation. *Proceedings of the Aristotelian Society* 48:111-138.
- Finlay, Stephen. 2007. Four faces of moral realism. *Philosophy Compass* 2 (6):820-849.
- FitzPatrick, William J. 2005. The Practical Turn in Ethical Theory: Korsgaard's Constructivism, Realism, and the Nature of Normativity. *Ethics* 115 (4):651-691.
- . 2008. Robust ethical realism, non-naturalism, and normativity. *Oxford studies in Metaethics*. Oxford: Oxford University Press. 3:159-205.
- . 2009. Recent work on ethical realism. *Analysis* 69 (4):746-760.
- Flew, Antony. 1986. *David Hume: philosopher of moral science*: Basil Blackwell Oxford.
- Foot, Philippa. 1958. Moral Beliefs. *Proceedings of the Aristotelian Society* 59:83-104.
- . 1972. Morality as a system of hypothetical imperatives. *The Philosophical Review* 81 (3):305-316.
- . 1978. Are Moral Considerations Overriding? In *Virtues and Vices*, edited by P. Foot. Los Angeles: University of California Press.
- Fotion, Nick. 2000. *John Searle, Philosophy Now*. New York: Acumen.
- Francén, Ragnar. 2010. Moral motivation pluralism. *The Journal of Ethics* 14 (2):117-148.
- Frankena, William K. 1939. The Naturalistic Fallacy. *Mind* 48 (192):464-477.

- Frege, Gottlob. 1892/1948. Sense and reference. *The Philosophical Review* 57 (3):209-230.
- Galdi, Silvia, Luciano Arcuri, and Bertram Gawronski. 2008. Automatic mental associations predict future choices of undecided decision-makers. *Science* 321 (5892):1100-1102.
- Gaus, Gerald F. 2011. *The order of public reason : a theory of freedom and morality in a diverse and bounded world*. New York: Cambridge University Press.
- Gert, Joshua. 2013. Overridingness, Moral. In *International Encyclopedia of Ethics*. New York: Blackwell
- Gibbard, Allan. 1992. *Wise Choices, Apt Feelings: A theory of normative judgment*: Boston: Harvard University Press.
- Glenn, A.L., R. Iyer, J. Graham, S. Koleva, and J. Haidt. 2009. Are all types of morality compromised in psychopathy? *Journal of personality disorders* 23 (4):384-398.
- Goodwin, Geoffrey P, and John M Darley. 2008. The psychology of meta-ethics: Exploring objectivism. *Cognition* 106 (3):1339-1366.
- Hare, R.M. 1952. *The Language of Morals* London: Oxford University Press, London.
- . 1954. Universalisability. *Proceedings of the Aristotelian Society* 55:295-312.
- . 1964. The promising game. *Revue Internationale de philosophie* 18:398-412.
- . 1981. *Moral thinking: its method, levels and point*. Oxford: Oxford University Press.
- Harman, Gilbert. 1975. Moral relativism defended. *The Philosophical Review* 84 (1):3-22.
- . 1977. Ethics and Observation. In *The Nature of Morality*. New York: Oxford University Press.
- . 1977. *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press.
- . 1982. Metaphysical Realism and Moral Relativism: Reflections on Hilary Putnam's Reason, Truth and History. *The Journal of Philosophy* 79 (10):568-575.
- . 1986. Moral Explanations of Natural Facts—Can Moral Claims Be Tested Against Moral Reality? *The Southern Journal of Philosophy* 24 (S1):57-68.
- Harman, Gilbert, and Judith Jarvis Thomson. 1996. *Moral Relativism and Moral Objectivity*. Oxford: Blackwell

- Hattiangadi, A. 2007. *Oughts and Thoughts: Rule-Following and the Normativity of Content*. New York: Clarendon Press.
- Hooker, Brad 2008. American Moral Philosophy. In *The Oxford Handbook of American Philosophy*, edited by C. Misak. New York: Oxford University Press.
- Horgan, Terence, and Mark Timmons. 1992. Troubles for new wave moral semantics: the 'open question argument' revived. *Philosophical Papers* 21 (3):153-175.
- . 1992. Troubles on Moral Twin Earth: moral queerness revived. *Synthese* 92 (2):221-260.
- Hume, David. 1739-40/2000. *A Treatise of Human Nature*, ed. David Fate Norton and Mary Norton. Oxford: Clarendon Press.
- Hussain, N., and N. Shah. 2006. Misunderstanding metaethics: Korsgaard's rejection of realism. *Oxford Studies in Metaethics* 1 Oxford: Oxford University Press:265-294.
- Hutchinson, Brian. 2001. *GE Moore's Ethical Theory*. New York: Cambridge University Press.
- Irwin, T. 2009. *The Development of Ethics, Volume 3: From Kant to Rawls*. Oxford: Oxford University Press.
- Ishikawa, S.S., A. Raine, T. Lencz, S. Bihrlé, and L. Lacasse. 2001. Autonomic stress reactivity and executive functions in successful and unsuccessful criminal psychopaths from the community. *Journal of Abnormal Psychology* 110 (3):423.
- Jackson, Frank. 1986. What Mary didn't know. *Journal of Philosophy* 83 (May):291-5.
- . 2000. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. New York: Clarendon Press.
- . 2012. In Defence Of Reductionism In Ethics. Unpublished Manuscript: <http://www.princeton.edu/~msmith/mycourses/JacksonReductionism2Mar2012.pdf>
- Jenkins, CS. 2005. Realism and independence. *American Philosophical Quarterly* 42 (3):199-209.
- Johansson, Petter, Lars Hall, Sverker Sikström, and Andreas Olsson. 2005. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* 310 (5745):116-119.
- Johansson, Petter, Lars Hall, Sverker Sikström, Betty Tärning, and Andreas Lind. 2006. How something can be said about telling more than we can know: On choice blindness and introspection. *Consciousness and Cognition* 15 (4):673-692.

- Joyce, Richard. 2007. *The Myth of Morality*. Cambridge: Cambridge University Press.
- Kant, I. 1785/2012. *Groundwork of the Metaphysics of Morals*. Edited by C. M. Korsgaard, M. Gregor and J. Timmermann. Cambridge: Cambridge University Press.
- Kiesewetter, Benjamin. 2012. A dilemma for Parfit's conception of normativity. *Analysis* 72 (3):466-474.
- Korsgaard, Christine. 2003. Realism and constructivism in twentieth-century moral philosophy. *Journal of Philosophical Research* 28:99-122.
- . 2009. *Self-constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- . 1996. *The Sources of Normativity*. New York: Cambridge University Press.
- Kramer, Matthew H. 2009. *Moral Realism as a Moral Doctrine, New Directions in Ethics*. Malden, MA: Wiley-Blackwell.
- Kripke, S.A. 1980. *Naming and Necessity*. Boston: Harvard University Press.
- Langford, C.H. 1968. Moore's Notion of Analysis. In *The Philosophy of G.E. Moore*, edited by P. A. Schilpp. Illinois: Open Court.
- Leiter, Brian. 2001. Moral facts and best explanations. *Social Philosophy and Policy* 18 (2):79-101.
- Lenman, J. 1999. The externalist and the amoralist. *Philosophia* 27 (3):441-457.
- Lewis, David. 1990. What Experience Teaches. In *Mind and Cognition*, edited by W. G. Lycan. Oxford: Blackwell.
- Lockie, Robert. 1998. What's wrong with moral internalism. *Ratio* 11 (1):14-36.
- Loeb, Don. 1998. Moral realism and the argument from disagreement. *Philosophical Studies* 90 (3):281-303.
- Mackie, John. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.
- Mandik, Pete. 2012. Color-Consciousness Conceptualism. *Consciousness and Cognition* 21 (2):617-631.
- McConnell, Terrance C. 1984. Objectivity and moral expertise. *Canadian Journal of Philosophy*:193-216.
- McDowell, John. 1988. Values and Secondary Qualities. In *Morality and Objectivity: A Tribute to J. L. Mackie*, edited by T. Honderich. New York: Routledge and Kegan Paul.

- McNaughton, D. 1988. *Moral Vision: An introduction to ethics*. Massachusetts:Wiley-Blackwell.
- Mele, Alfred R. 1989. Motivational internalism: The powers and limits of practical reasoning. *Philosophia* 19 (4):417-436.
- . 1996. Internalist moral cognitivism and listlessness. *Ethics* 106 (4):727-753.
- Mendelsohn, R.L. 2005. *The Philosophy of Gottlob Frege*. Cambridge: Cambridge University Press.
- Mill, John Stuart. 1863. *Utilitarianism*. London: Parker, Son, and Bourn.
- Miller, Alexander. 2003. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press.
- Moore, George Edward. 1903/1993. *Principia Ethica*. Cambridge: Cambridge University Press.
- Morgan, Seiriol. 2006. Naturalism and Normativity. *Philosophy and Phenomenological Research* 72 (2):319-344.
- Nagel, Thomas. 1970. *The Possibility of Altruism*. Oxford: Oxford Clarendon Press.
- . 1996. Universality and the reflective self. In *The Sources of Normativity*, edited by O. O'Neill. Cambridge: Cambridge University Press.
- Nichols, Shaun. 2002. Norms with feeling: Towards a psychological account of moral judgment. *Cognition* 84 (2):221-236.
- . 2004. After objectivity: An empirical study of moral judgment. *Philosophical Psychology* 17 (1):3-26.
- Nisbett, Richard E, and Nancy Bellows. 1977. Verbal reports about causal influences on social judgments: Private access versus public theories. *Journal of Personality and Social Psychology* 35 (9):613.
- Nisbett, Richard E, and Timothy D. Wilson. 1977. Telling More than We Can Know: Verbal reports on mental processes. *Psychological Review* 84 (3):231.
- Nowell-Smith, PH. 1952. Ethics (London, 1954). *The word "philosopher" is rarely, if ever, used except in this patronising sense, and the usage is common among present-day holders of Chairs in philosophy.*
- Nuccetelli, S., and G. Seay. 2013. Does Analytical Moral Naturalism Rest on a Mistake? In *Ethical Naturalism: Current Debates*, edited by S. Nuccetelli and G. Seay. Cambridge: Cambridge University Press.
- Olson, Jonas. 2002. Are desires de dicto fetishistic? *Inquiry* 45 (1):89-96.

- Papineau, David. 2006. Seven Phenomenal and Perceptual Concepts. *Phenomenal concepts and phenomenal knowledge: New essays on consciousness and physicalism*. Edited by Torin Alter and Sven Walter. Oxford: Oxford University Press..
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- . 2011. *On What Matters*. Vol. 1. Oxford: Oxford University Press.
- . 2011. *On What Matters*. Vol. 2. Oxford: Oxford University Press.
- Pettit, Philip. 2001. Embracing Objectivity in Ethics. In *Objectivity in Law and Morals*, edited by B. Leiter. Cambridge: Cambridge University Press.
- Pigden, Charles R. 1989. Logic and the Autonomy of Ethics. *Australasian Journal of Philosophy* 67 (2):127-151.
- . 1991. Naturalism. In *A Companion to Ethics: Blackwell Companions to Philosophy*, edited by P. Singer. Oxford: Blackwell
- . 2010. *Hume on Is and Ought*. Indiana: Palgrave Macmillan.
- . 2013. Is–Ought Gap. In *The International Encyclopedia of Ethics*. New York: Blackwell.
- Plato. 2002. Five Dialogues: Euthyphro, Apology, Crito, Meno, Phaedo. edited by G. M. A. Grube. Indianapolis: Hackett Publishing Company.
- Platts, Mark. 1980. Moral Reality and the End of Desire. In *Reference, Truth and Reality: Essays in the Philosophy of Language*, edited by M. Platts. London: Routledge and Keagan Paul.
- Prichard, H.A. 1912. Does moral philosophy rest on a mistake? *Mind*:21-37.
- Prinz, Jesse. 2006. The emotional basis of moral judgments. *Philosophical Explorations* 9 (1):29-43.
- . 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.
- Prior, Arthur N. 1960. The autonomy of ethics. *Australasian Journal of Philosophy* 38 (3):199-206.
- Pronin, Emily, and Matthew B Kugler. 2007. Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology* 43 (4):565-578.
- Putnam, Hilary. 1992. *Realism with a Human Face*. Boston:Harvard University Press.

- Rachels, James. 1998. Introduction. In *Ethical Theory: The Question of Objectivity*. Oxford: Oxford University Press.
- . 2007. The challenge of cultural relativism. In *Bioethics: an Introduction to the History, Methods, and Practice*, edited by A. R. J. Nancy S. Jecker, Robert A. Pearlman. Massachusetts: Jones and Bartlett Publishers.
- Radzik, L. 1999. A normative regress problem. *American Philosophical Quarterly* 36 (1):35-47.
- Railton, P. 2009. Internalism for externalists. *Philosophical Issues* 19 (1):166-181.
- Railton, Peter. 1986. Moral realism. *The Philosophical Review* 95 (2):163-207.
- . 2013. Two Sides of the Meta-Ethical Mountain? In *Does Anything Really Matter? Responses to Parfit/Parfit on Objectivity*, edited by P. Singer. Forthcoming.
- Rasmussen, Stig Alstrup. 1985. Quasi-realism and mind-dependence. *The Philosophical Quarterly* 35 (139):185-191.
- Rosenthal, David M. 2001. Color, mental location, and the visual field. *Consciousness and Cognition* 10 (1):85-93.
- . 2010. How to think about mental qualities. *Philosophical Issues* 20 (1):368-393.
- Roskies, Adina. 2003. Are ethical judgments intrinsically motivational? Lessons from "acquired sociopathy"[1]. *Philosophical Psychology* 16 (1):51-66.
- . 2006. Patients with ventromedial frontal damage have moral beliefs. *Philosophical Psychology* 19 (5):617-627.
- . 2008. Internalism and the Evidence from Pathology. *Moral psychology* 3:191-206.
- Ross, W.D. 1939. *Foundations of Ethics: The Gifford Lectures*. Oxford: Clarendon Press.
- Ryle, Gilbert. 2009. *The Concept of Mind: 60th Anniversary Edition*. New York: Taylor & Francis.
- Sadler, B.J. 2003. The possibility of amorality: A defence against internalism. *Philosophy* 78 (1):63-78.
- Sadler, Brook Jenkins. 2000. Can the amoralist only be 'right'? A closer look at the inverted-commas argument. *Southwest Philosophy Review* 17 (1):113-122.
- Sandberg, Joakim, and Niklas Juth. 2011. Ethics and Intuitions: A Reply to Singer. *The Journal of Ethics* 15 (3):209-226.

- Sarkissian, Hagop, John Park, David Tien, Jennifer Cole Wright, and Joshua Knobe. 2011. Folk moral relativism. *Mind & Language* 26 (4):482-505.
- Sayre-McCord, Geoffrey. 1986. The Many Moral Realisms. *The Southern Journal of Philosophy* 24 (S1):1-22.
- . 1997. The Metaethical Problem. *Ethics* 108 (1):55-83.
- Schlick, Moritz. 1939. *Problems of Ethics*. New York: Prentice Hall Incorporated.
- Schmidt, Thomas. 2012. Moral values and the fabric of the world. In *Normativity and Naturalism*, edited by P. Schaber: BoD—Books on Demand.
- Schroeder, Mark. 2011. What Matters about Metaethics? In *Does Anything Really Matter? Responses to Parfit/Parfit on Objectivity*, edited by P. Singer: University of Southern California. Forthcoming.
- Searle, John R. 1964. How to derive "ought" from "is". *The Philosophical Review* 73 (1):43-58.
- . 1979/1985. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Setiya, Kieran. 2011. On What Matters, by Derek Parfit. *Mind* 120 (480):1281-1288.
- Shafer-Landau, Russ., and Terence. Cuneo. 2006. *Foundations of Ethics: An Anthology*. Massachusetts: Wiley.
- Shafer-Landau, Russ. 2000. A defense of motivational externalism. *Philosophical Studies* 97 (3):267-291.
- . 1994. Ethical disagreement, ethical objectivism and moral indeterminacy. *Philosophy and Phenomenological Research* 54 (2):331-344.
- . 1994. Supervenience and Moral Realism. *Ratio* 7 (2):145-152.
- . 2003. *Moral realism: A Defence*. Oxford: Oxford University Press
- . 2012. Moral Rationalism. In *Ethical Theory: An Anthology*, edited by R. Shafer-Landau. New York: Wiley.
- Shepski, Lee. 2008. The vanishing argument from queerness. *Australasian Journal of Philosophy* 86 (3):371-387.
- Silverstein, Matthew. 2012. Inescapability and Normativity. *Journal of Ethics and Social Philosophy* 6 (3).
- Singer, Peter. 1995. *Animal Liberation*. New York: Random House.

- . 2005. Ethics and Intuitions. *The Journal of Ethics* 9 (3-4):331-352.
- Sinnott-Armstrong, Walter. 2007. *Moral Skepticism*. Oxford: Oxford University Press.
- . 2010. Mackie's Internalisms. In *A World Without Values: Essays on John Mackie's Moral Error Theory*, edited by R. Joyce and S. Kirchin. New York: Springer.
- Smith, Michael. 1995. *The Moral Problem*. Oxford: Blackwell
- Snare, Francis. 2002. *Morals, Motivation, and Convention: Hume's Influential Doctrines*. Cambridge: Cambridge University Press.
- Soames, Scott. 2003. *Philosophical Analysis in the Twentieth Century*. 2 vols. Princeton, N.J.: Princeton University Press.
- Stern, R. 2010. Moral scepticism and agency: Kant and Korsgaard. *Ratio* 23 (4):453-474.
- Stevenson, Charles Leslie. 1937. The emotive meaning of ethical terms. *Mind* 46 (181):14-31.
- Stocker, Michael. 1979. Desiring the bad: An essay in moral psychology. *The Journal of Philosophy* 76 (12):738-753.
- Stoljar, Daniel. 2005. Physicalism and phenomenal concepts. *Mind and Language* 20 (2):296-302.
- Strandberg, Caj. 2004. In defence of the open question argument. *The Journal of ethics* 8 (2):179-196.
- . 2007. Externalism and the content of moral motivation. *Philosophia* 35 (2):249-260.
- Strout, Rowland 2010. Twentieth Century Moral Philosophy In *The Routledge Companion to Twentieth Century Philosophy*, edited by D. Moran. New York: Routledge.
- Sturgeon, Nicholas. 1985. Moral explanations. *Morality, Reason, and Truth*, edited by D. Copp and D. Zimmerman. Totowa, New Jersey: Rowman & Littlefield. 49-78.
- Svavarsdóttir, Sigrún. 1999. Moral cognitivism and motivation. *The Philosophical Review* 108 (2):161-219.
- . 2005. How do moral judgments motivate. In *Contemporary Debates in Moral Theory*, edited by J. Dreier. Oxford: Wiley-Blackwell.
- Tännsjö, Torbjörn. 2010. *From reasons to norms: on the basic question in ethics*. Vol. 22. New York: Springer.

- Thomson, Judith Jarvis. 2003. The Legacy of Principia. *The Southern Journal of Philosophy* 41 (S1):62-82.
- Timmons, Mark. 1998. The Logic of Ethical Discourse. In *Routledge Encyclopedia of Philosophy*. New York: Routledge.
- Tye, Michael. 2002. *Consciousness, color, and content*. Massachusetts: The MIT Press.
- Umphrey, Stewart. 2002. *Complexity and analysis*. Maryland: Lexington Books.
- Unger, Peter. 1966. On experience and the development of the understanding. *American Philosophical Quarterly* 3 (1):48-56.
- Vasiliou, Iakovos. 2008. *Aiming at Virtue in Plato*. Cambridge: Cambridge University Press.
- Watkins, Eric, and William J. FitzPatrick. 2002. O'Neill and Korsgaard on the Construction of Normativity. *The Journal of Value Inquiry* 36 (2):349-367.
- Wiggins, David. 1990. Moral cognitivism, moral relativism and motivating moral beliefs. *Proceedings of the Aristotelian Society*.
- Williams, Bernard. 1980. Internal and external reasons. In *Moral Luck*, edited by B. Williams. Cambridge: Cambridge University Press.
- Wilson, Timothy D. 2009. Know thyself. *Perspectives on Psychological Science* 4 (4):384-389.
- Wilson, Timothy D., and Yoav Bar-Anan. 2008. The unseen mind. *Science* 321 (5892):1046-1047.
- Wong, David. 1991. Relativism. In *A Companion to Ethics*, edited by P. Singer. Oxford: Blackwell.
- Zangwill, Nick. 1990. Quasi-quasi-realism. *Philosophy and Phenomenological Research* 50 (3):583-594.
- . 1994. Moral mind-independence. *Australasian Journal of Philosophy* 72 (2):205-219.
- . 2003. Externalist moral motivation. *American Philosophical Quarterly* 40 (2):143-154.
- . 2008. The indifference argument. *Philosophical Studies* 138 (1):91-124.
- Zimmerman, David. 1980. Meta-ethics naturalized. *Canadian Journal of Philosophy*:637-662.