

MORAL OBJECTIVITY, AUTONOMY, AND REASONS:
THE CONSTRUCTIVIST CHALLENGE TO REALISM

A Dissertation

Presented to the Faculty of the Graduate School
of Cornell University

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by

Nathaniel Nelson Jezi

May 2011

© 2011 Nathaniel Nelson Jezzi

MORAL OBJECTIVITY, AUTONOMY, AND REASONS:
THE CONSTRUCTIVIST CHALLENGE TO REALISM

Nathaniel Nelson Jezzi, Ph. D.

Cornell University 2011

This dissertation defends realist views in ethics against arguments advanced by ethical constructivists. Realists think that ethical truth obtains independently of any preferred perspective. That is, it is at bottom independent of whatever beliefs, desires, or other forms of commitment we have. Constructivists, in contrast, deny this. But this could mean different things, and there is currently no consensus on what constructivism involves exactly or how we ought to understand this dispute. Consequently, it has been difficult to evaluate the merits of constructivist arguments. This work attempts to remedy these deficiencies by developing a framework for both interpreting and evaluating the nature and scope of the constructivist's challenge.

In the first two chapters, I address some of the larger interpretative disputes. In Chapter 1, I argue that there are two main ways of understanding constructivism. Each of these corresponds to the rejection of a particular commitment of realism. In Chapter 2, however, I argue that neither of these should be understood as representing a free-standing view in metaethics. Rather, each takes aim at a narrower target: viz., the realist's conception of ethical *objectivity*. The first type of constructivist challenge rejects realist claims about the *nature* of ethical objectivity; the second type accepts

these but rejects claims about its *scope*.

In the final three chapters, I evaluate these two versions of the challenge. In Chapter 3, I argue that if the constructivist rejects *all* stance-independent ethical truths, she commits herself to absurd results within ethical theory. This prompts me to consider more modest constructivist theses, ones that allow for *some* stance-independent ethical truths but that also significantly restrict the scope that such truths play within an ethical theory. In Chapters 4 and 5, however, I argue that the best arguments for this more modest constructivism also fail because the ethical considerations they appeal to – i.e., moral rationalism and autonomy – can be equally accommodated by a robust moral realism.

BIOGRAPHICAL SKETCH

Nathaniel Jezzi received his B.A. in Government with a concentration in Political Philosophy from Georgetown University in 1998. From 2000-2003, he studied Philosophy, Sociology, and Greek Philology in Berlin, Germany at both the Humboldt Universität and Freie Universität. In 2003, he began his graduate studies in the Sage School of Philosophy at Cornell University. He received his M.A. in Philosophy from Cornell University in 2007. He has taught at Cornell University, the Elmira Correctional Facility, and the University of Aberdeen. He currently lives in Scotland.

For Katie

ACKNOWLEDGMENTS

I would like to express my deep gratitude to all of the people who have helped with this project. I am particularly indebted to my advisor Nicholas Sturgeon, not only for his many probing questions and helpful comments, but also for his patience and generosity. His rigor, clarity, and integrity of thought have been a model for me. I have surely failed to match these high standards. But, to the extent that this work shares in these qualities, it is in large part due to his influence. I am also very much indebted to the other three members of my committee – Terence Irwin, Matti Eklund, and Michelle Kosch – for their support and advice. The idea for this project developed during my third-year tutorial with Terry. He allowed me the space to find my own questions and gave me the confidence to pursue them. Matti commented on every draft along the way and pushed me to think harder and deeper on a host of topics. Michelle was the last to join the team. She has been a booster, as well as an honest and incisive critic. I owe her, along with the rest of my committee, so much more than I am able to express here.

I have also benefited incredibly from my friends and fellow graduate students at Cornell. I am especially thankful for the comments and support of Jacob Klein, Raúl Saucedo, Daniel Koltonski, and Patrick Mayer. They read and commented on multiple drafts, often on very short notice, and offered much needed encouragement. I thank each of them for their patience and friendship.

I would like to thank the Sage School of Philosophy for giving me the opportunity to pursue my graduate studies. I may not have been the most deserving of candidates, but I have certainly benefited from their generosity. The intellectual and social milieu of the Department enriched my life greatly.

And, of course, I owe no small thanks to my family. My parents and two

brothers have been a constant source of support and inspiration. My wife, however, deserves more credit than anyone else for helping me across the finish line. She has been my closest friend and sharpest critic. It is to her that I devote this work.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vii
1. The Constructivist's Challenge	1
1.1 Introduction	1
1.2 Understanding Constructivism Against the Backdrop of Realism	6
1.3 Naturalistic Motivations	12
1.4 Motivations From Within Ethical Theory	15
1.5 Preserving the Challenge of Substantive Constructivism	18
1.6 Offending Realist Intuitions	28
1.7 The General Argumentative Strategy	31
2. Is Constructivism a <i>Free-Standing</i> Metaethical View?	34
2.1 Introduction	34
2.2 The Proceduralist Characterization	36
2.3 Constructivism as a Response Dependence View	39
2.4 The Standpoint Characterization	43
2.5 Problems for the Standpoint Characterization	52
2.6 Reframing the Challenge	61
3. Formal Constructivism and Considered Moral Judgments	64
3.1 Introduction	64
3.2. What is Formal Constructivism?	66
3.3 An Internal Challenge for Formal Constructivism	67
3.4 Two Kinds of Formal Constructivism	70

3.5	The Failure of Formal Kantian Views	76
3.6	Is the Challenge Too Strong?	84
3.7	What if the Constructivist Rejects Moral Rationalism?	87
3.8	Conclusion	94
4.	Normative Alienation and the Morality of Common Sense	98
4.1	Introduction	98
4.2	The Explanatory Challenge	102
4.3	The Argument from Moral Reasons	107
4.4	Rational Force: <i>A Second Pass</i>	114
4.5	Rational Force: <i>A Third Pass</i>	119
4.6	Rational Force: <i>A Last Pass</i>	124
4.7	Objection: Deliberation Without Construction	128
4.8	Conclusion	131
5.	Autonomy and the Nature of Morality	132
5.1	Introduction	132
5.2	Some Preliminaries	137
5.3	What is Korsgaard's Argument for SL?	139
5.4	My Strategy for Evaluating the Argument	142
5.5	Reconstructing Korsgaard's Argument	143
5.6	Moral Agency and Responsibility	150
5.7	Two Incompatible Senses of Autonomy	158
5.8	Why Korsgaard's Response Fails	161
5.9	Conclusion	164
	Works Cited	168

CHAPTER 1

THE CONSTRUCTIVIST'S CHALLENGE

1.1. Introduction

Human beings do all kinds of things. While some of these activities are trivial or accidental to our nature, others are central to our understanding of what it is to be human. Ethical practice is arguably one of these. As ethical creatures, we engage in certain characteristic activities: *We value things. We take some actions to be wrong and some to be right. We also take some things to count as reasons for acting.* There may be much debate about what the proper objects of these activities are, but it is *not* controversial that we engage the world in these ways. Human beings might have evolved to value quite different things or take different kinds of actions to be wrong or right or take slightly different features of the world to count as reasons for action. But a mature, healthy human being who failed to engage in any of these characteristic activities in any way would be a very strange creature indeed. The central question of this dissertation will be how we are to understand the relation between these activities and the nature of what ethical terms and expressions stand for.¹

Although contemporary moral philosophy has generated a flourish of literature on this topic recently, the question is a very old one. It is arguably a version of the one that Plato entertains in his presentation of a dialogue between Socrates and

¹ Here and throughout I use “ethical” to refer to a broad range of topics that are the concern of practical philosophy – including morality, value, and practical reason.

Euthyphro.² There the topic is *piety*. Specifically, is something pious because the gods love it? Or do the gods love it because it is pious? Plato's brief arguments in the dialogue have been interpreted in different ways and done little to quiet interest in these questions. In order to broaden the terms of the discussion, we may recast Euthyphro's question in a "rough, secular paraphrase".³

- (i) Do we value things because they are valuable? Or do things have value because we value them?
- (ii) Do we take certain actions to be wrong (or right) because they are wrong (or right)? Or are these actions wrong (or right) because we take them to be so?
- (iii) Do we take certain features of the world to count as reasons for action because they are reasons? Or do these features count as reasons for action because we take them to be so?

These are metaphysical questions. How we answer them will say something about the kinds of facts or properties that exist and what they are like – for example, whether an account of these facts and properties must make essential reference to these activities or the standpoints that characterize them.⁴ *Ethical constructivists* answer "no" to the first of these pairs of questions and "yes" to the second. *Ethical Realists* answer "yes" to the first question and "no" to the second.⁵

² Plato (1997), 1-16.

³ This expression comes from Street (2010), 370.

⁴ What exactly a *standpoint* is will be discussed a bit later.

⁵ Things are actually more complicated here. Neither constructivism nor realism may be alone in taking these particular stances, depending on how one defines each view. For example, one might object that

Although these questions help to provide an intuitive grasp of what is at issue, they do not shed any real light on what distinguishes constructivism from realism. My task in this dissertation will be to explore different ways of understanding constructivism as well as the arguments that support them.

Many philosophers have shown an interest in these questions – either presenting (Onora O’Neill, Stephen Darwall, Allan Gibbard, and Peter Railton), or defending (John Rawls, Ronald Milo, Christine Korsgaard, T.M. Scanlon, Sharon Street, Aaron James), or critiquing (David Brink, Mark Timmons, Allan Gibbard, William Fitzpatrick, Nadeem Hussain and Nishi Shah, Philip Pettit, Russ Shafer-Landau, David Enoch, Michael Ridge) versions of constructivism.⁶ But there is currently no consensus on what constructivism about any of these topics is supposed to be or how exactly it differs from realism or other available metaethical positions, like response-dependence theories or expressivism.⁷ This lack of consensus is in part spurred by the fact that one may be a constructivist about one area of ethical discourse

the distinction is not merely between constructivists and realists but between constructivists and realists *or* quasi-realists. Defenders of quasi-realism, like Simon Blackburn (1984), explicitly claim that they side with realists on this issue. However, it is controversial whether they are entitled to this claim (in Blackburn’s terms, whether they have really *earned the right* to such a response). Similarly, one may object that a response-dependence theorist or a speaker subjectivist, more generally, would respond to these questions in the same way as the constructivist does. I will have more to say about how constructivism contrasts with these views in the next chapter. My claim here is just that these questions provide one helpful way of highlighting what is at issue in debates between constructivists and realists, even if this distinction captures broader families of views under which constructivism and realism each respectively fall.

⁶ See O’Neill (2003), Darwall, Gibbard, Railton (1992), Rawls (1980), Milo (1995), Korsgaard (1996), (2003), (2009), Scanlon (1982), (1998), (2003), Street (2008), (2009), (2010), James (2007), Brink (1989), Timmons (2003), Gibbard (1999), Fitzpatrick (2005), Hussain and Shah (2006), Parfit (forthcoming), Pettit (2000), Shafer-Landau (2003), Enoch (2009), Ridge (forthcoming).

⁷ Street (2010):

It should be emphasized at the outset that the question -- ‘What is constructivism in ethics and metaethics?’ -- has no uncontroversial answer at the present time. (p.364)

while maintaining a realist, or expressivist, etc., view about others.⁸ For example, one may be a constructivist about morality but a realist about practical reason,⁹ or a constructivist about value but a realist about morality, or a constructivist about practical reason and a realist about value and morality. But one may also be a constructivist, or realist, about all three.¹⁰ It may turn out that there are different motivations and arguments for constructivism depending on whether one is concerned with value, morality, or practical reason. What's more, there may turn out to be some differences in the metaphysics depending on which area is at issue.

In order to set aside some of these complications, I will take as my target *ethical constructivism* most generally. Whatever differences in motivation and argument one may find for each of these specific views, there is also a lot that each view has in common. For the sake of simplicity, I will assume that an evaluation of

⁸ Each of the three positions is independent of the other two. But this independence is won at the price of rejecting what are often taken to be plausible views about the relations between reasons, morality, and value. For example, many people think that there is some internal connection between having a reason to perform some action and that action's moral or evaluative status, such that an action's rightness *necessarily* provides some reason to perform it or that an object's or state-of-affair's goodness *necessarily* provides some reason to pursue or promote it. But if this is true, it might appear as if our best story about what a *reason* is will take the same form as our best story about what *rightness* or *goodness* is. If reasons are constructed, then so is the right, or the good. Otherwise, it would appear difficult to explain the *necessity* of the above relationship. For, as Humeans are wont to point out, it is difficult to explain a necessary connection between distinct existences. However, if one rejects the claim that these things are necessarily related there is no pressure to account for them in similar terms. Rejecting the thesis arguably bears some costs. But it illustrates how accounts of reasons might be thought to come apart from accounts of morality and value. Of course, this does not explain how constructivism about value and morality are independent from each other. In this case, the relevant view that must be rejected is consequentialism. Consequentialists think that moral claims reduce in some way to claims about value. If this is right, it follows that a constructivism about value entails constructivism about morality. However, if one rejects consequentialism, there is little pressure to account for morality and value in similar terms. Depending on one's views about the plausibility of consequentialism, however, rejecting the view might involve a significant cost. Nevertheless, one can see how constructivism about value might come apart from constructivism about morals.

⁹ Scanlon (1998).

¹⁰ Those who claim to defend a comprehensive constructivism include, e.g., Korsgaard (1996) and Street (2008) [NB Street appears to have rejected constructivism about morality in more recent work; see Street (2010)]. Comprehensive realists include, e.g., Enoch (2007), Fitzpatrick (2008), Parfit (forthcoming), Shafer-Landau (2003).

this common ground will suffice for evaluating the plausibility of *any* version of constructivism. In other words, if the general arguments for constructivism succeed, this should provide good reason for thinking that a particular version of constructivism succeeds. Alternatively, if these general arguments fail, this should provide good reason for discounting versions of these arguments that support a particular type of constructivism.¹¹ This assumption finds support across the constructivist literature and is borne out by (what I will show to be) the general nature of constructivist arguments. Although I will sometimes take issue with a particular form of constructivism, this is only because I am concerned to respond to the specific terms of the arguments I am addressing. Nevertheless, my target throughout should be understood as ethical constructivism most generally.

In this first chapter, my aim is to present some of the basic motivations for constructivism, as well as my own motivation for evaluating the view, and outline the constructivist's basic argumentative strategy. This approach will draw our attention to the central themes and questions I will address in the course of this work as well as give us a sense for the structure of the discussion that follows in subsequent chapters.

¹¹ In other words, the positions themselves are logically independent of one another, but the arguments that support one version or another supports them all. This is because the arguments are general. They do not discriminate between value, morality, or practical reason.

1.2 Understanding Constructivism Against The Backdrop of Realism

One standard way of framing constructivism is to present it as a response to the putative failings of ethical realism.¹² But how we are to understand realism, or anti-realism, in ethics is itself contested. This, of course, complicates things. If constructivism is presented as a response to realism, but the commitments of realism are themselves contested, it would not appear as if we are starting with the clearest framework for understanding constructivism. Despite this complication, however, I think the framework can be a useful one. The situation just requires that we be explicit about how we are to understand the term *realism* in this context.

It is fairly uncontroversial to take *ethical realism* to include at least the following two conditions:

- (1) Atomic ethical statements are the kinds of things that may be literally true or false.
- (2) At least some of them, literally construed, are true.¹³

¹² This is certainly not an uncontroversial way of framing the view. Some might argue that this gets things exactly backwards. For example, Korsgaard (1996), and especially (2003) – see p. 101 – argues that we should understand realism in moral philosophy as a series of historical reactions to every ambitious positive view about the nature of morality that has been proposed (including, e.g., constructivism and its close relatives). This historical claim is extremely contentious. But evaluation of it is beyond the scope of this work. Regardless of whether realism has been a reactive view or the default position in the history of moral philosophy, my claim is that we may still learn something about what constructivism is and why its proponents defend it by contrasting it with the commitments and putative failings of realism.

¹³ See, e.g., Sayre-McCord (1988). NB It is interesting that this kind of definition, quite controversially, ends up counting certain views as versions of realism – e.g., those advanced respectively by Foot (1983), Railton (1986), and Smith (1989), (1993) – despite the fact that they are not always categorized this way.

These conditions look promising insofar as they serve to contrast realism with two commonly recognized anti-realist competitors. The first condition contrasts the view with *non-cognitivism* or *expressivism*.¹⁴ Defenders of these views deny that ethical statements are straightforwardly or literally fact stating; rather, they claim that we use ethical language to express some non-belief-like state of mind. The second condition contrasts the view with an *error theory*. Defenders of this kind of view accept that we use ethical language to report beliefs but claim that all of these beliefs are systematically false because ethical terms and expressions fail to refer to anything. Again, both types of views are framed in opposition to realism. Insofar as (1) and (2) achieve this contrast, they provide a helpful way of understanding realism's commitments.

These two conditions also appear sufficient to distinguish realism from *some* statements of constructivism. For example, Christine Korsgaard has described constructivism in ways that look incompatible with (1).¹⁵ On her account, *atomic* ethical statements are not made true by facts that we may come to know and apply in deliberation; rather, they express practical problems that agents must solve. The details of this proposal are not completely clear, but some have argued that Korsgaard's view

¹⁴ Some philosophers would claim that these two positions amount to the same thing, or perhaps that the latter is species of the former. Other philosophers have voiced objection to such conflation. For example, Simon Blackburn (1984) defends expressivism but rejects the label "non-cognitivism" for his view. This is because this label is usually applied to views that rule out the possibility of moral beliefs and moral knowledge. However, Blackburn argues that expressivists can *earn the right* to talk about moral belief and knowledge. The strategy of accommodating this kind of realist talk within a non-realist metaphysical framework is what he calls "quasi-realism".

¹⁵ See Korsgaard (2003). In her most recent work, Sharon Street also describes constructivism in ways that make it unclear whether the view satisfies (1). See, especially, her discussion of the similarities between constructivism and expressivism in (2010), 375ff. I will have more to say about Street's proposal later in the discussion.

does not construe moral truth literally.¹⁶ If this were indeed the case, the first two conditions alone would suffice for distinguishing realism from constructivism. But even if some statements of constructivism might be ruled out by these conditions, others are not. In fact, part of what many take to be attractive about constructivism is that it *does* satisfy these two conditions. By taking on board some of the features of realism and rejecting others, constructivists claim to capture all that is attractive about realism while avoiding standard objections against it. If constructivism failed to satisfy (1) or (2), a defender of the view could not claim any such advantage – for without them there is nothing of realism to speak of. This means that we need to add some other condition(s) to our account of realism, one or another that captures the distinction that these other constructivists have in mind.

Russ Shafer-Landau has proposed the following candidate:

- (3) There are moral truths that obtain independently of any preferred perspective, in the sense that *the moral standards that fix the moral facts are not made true by virtue of their ratification from within any actual or hypothetical perspective.*¹⁷

¹⁶ See Fitzpatrick (2008), 161-2 for discussion.

¹⁷ Shafer-Landau (2003), 15. In his characterization of a *robust ethical realism*, Fitzpatrick (2008) averts to this condition and later proposes his own modification (p. 166).

This condition is used to describe what is sometimes referred to as the *stance-independence* of ethical facts and properties.¹⁸ This is because it makes at least some instances of moral truth independent of “any preferred perspective”, actual or hypothetical. A *perspective*, or *standpoint* (as I will later refer to it), is a complex system of intentional psychological states (i.e., stances), such as beliefs, desires, commitments, reactive attitudes, etc.

According to (3), even if some ethical standards come into existence because they figure as the objects of our desires, or choices, or beliefs, etc. (e.g., traffic laws, rules of etiquette), there are some that exist independently of our intentional stances. In other words, these standards would exist even if we (or idealized versions of ourselves) desired, wished, or believed that they did not, or decided to ignore them. I will soon have occasion to say more about what a perspective, or standpoint, is. For now, however, it is just important to note how this third condition serves to contrast realism with constructivism. Unlike the first two conditions, this one does appear to get at the distinction many constructivists have in mind.

Specifically, it would appear to give some substance to the intuitive sense of dependence implicit in the questions stated earlier. These questions suggest that constructivism differs from realism insofar as it makes ethical facts and properties depend on our ethical practice in some essential way. However, it is important to note that this third condition does not entirely rule out the characterization of ethical facts

¹⁸ For coinage of this expression, see Milo (1995), 182, 192. Both Shafer-Landau (2003) and Street (2008) follow Milo in describing realism as committed to a stance-independent view of moral truth. Fitzpatrick (2008) introduces this term in discussing Shafer-Landau’s view.

and properties in terms of psychological facts about the agent. For there are different ways in which ethical facts and properties might depend on an agent's psychology.

Let us distinguish two types of mind dependence.¹⁹ On the one hand, let us say that ethical facts and properties are *weakly mind dependent* if they are constituted by mental states in some essential way. For example, a hedonistic utilitarianism that characterizes ethical facts and properties in terms of an agent's experiences of pleasure and pain might count as mind dependent in just this way. On the other hand, let us say that ethical facts and properties are *strongly mind dependent* if their ethical status depends in some essential way on their being made the object of an agent's intentional psychological state (e.g., a belief, desire, wish, etc.). In short: this distinction does not concern whether ethical facts and properties are characterized psychologically – since both views allow for this much. Rather, it concerns *how* the status of these facts and properties gets routed through an agent's psychology.

Importantly, the stance-independence condition in (3) only rules out that ethical facts and properties are strongly mind dependent; it is compatible with their being weakly mind dependent. The reason for this is that realists are keen to preserve a particular, and arguably intuitive, sense of *objectivity* – one according to which the matter of ethical truth is not “up to us”. If ethical facts and properties were strongly mind dependent, the truth of ethical statements would be up to us in the sense that we would make them true on the basis of our desiring, believing, or willing it to be the case. By contrast, if the truth-making facts are merely weakly mind dependent, this is not so. Although a hedonistic utilitarianism, for example, makes ethical truth depend

¹⁹ Here I am closely following the distinction introduced by Milo (1995), 191.

on facts about an agent's psychology, it does not leave it up to the agent (actual or hypothetical) to make it the case that a particular ethical truth obtains.

If we accept that realism offers a stance-independent view of ethical facts and properties, constructivism, by contrast, ought to be understood as a species of a *stance-dependent* view. On this account, there are *no* moral, or ethical, truths that obtain entirely independently of any preferred perspective. The standards that fix the relevant class of ethical facts are *always* made true by virtue of their ratification from within some actual or hypothetical perspective. Constructivists offer various characterizations of the relevant status-conferring perspective or standpoint, as well as the kind of ratification that is required. Some of these differences will be discussed shortly; others will arise in the course of discussions in subsequent chapters.

Many constructivists are happy to accept (1) and (2), but they argue that realists go too far in positing (3). Again, this condition constitutes, in large part, the realist notion of ethical objectivity. For this reason, we might take constructivists to be rejecting the idea that ethical facts and properties are objective in the realist's sense – while leaving open the possibility that they might count as objective in some other sense. Some constructivists argue that, by incorporating (3), realists fail to accommodate deeply held philosophical commitments.²⁰

²⁰ But, as I will argue later, not *all* constructivists reject (3). See discussion in sections 4-6.

1.3 Naturalistic Motivations

The objection is that realism fails to accommodate our broader metaphysical and epistemological commitments. Here the concern is generally that realism about value, or morality, or practical reason is incompatible with *philosophical naturalism*. Very roughly, this is the view that the only kinds of facts and properties that exist are *natural* ones – i.e., those facts and properties that (could) figure as the objects of investigation of our best scientific practices. The alleged problem is that ethical facts and properties could only satisfy condition (3) if naturalism were false.

There are different versions of this argument in the literature. According to one popular version of the objection, ethical facts and properties exhibit certain regular connections with our motivational capacities, such that, for example, anyone acquainted with an action's wrongness is thereby motivated to avoid it. If these motivational connections are understood naturalistically (e.g., as necessary connections between ethical properties and an agent's desires or dispositions to choose), it is hard to see how ethical facts and properties could enjoy the independence described in condition (3). They would have to be stance-independent by nature yet necessarily connected with certain motivational stances. The worry is that this would suggest, in the words of John Mackie, that ethical facts and properties were "utterly different from anything else in the universe."²¹ The conclusion here is that realism commits one to a kind of metaphysical *queerness*.

²¹ Mackie (1977), 38.

The claim of metaphysical queerness gives rise to a related concern about epistemological queerness. Recently, Sharon Street has advanced a version of this kind of naturalistic argument.²² She claims that ethical realism is incompatible with our best evolutionary account of how we came to make the ethical judgments we do.

According to this argument, if realism were true, we would have no good explanation for how our ethical judgments have succeeded in matching (or “tracking”) stance-independent ethical truths; rather, the truth of these judgments would have to be entirely a matter of unlikely coincidence.

Constructivism, by contrast, is supposed to avoid these problems. By grounding ethical truths in features of intentional states, constructivists claim that their views only makes use of naturalistic materials, e.g., ones that can be accounted for by empirical psychology. These are features that may be appealed to in order to explain the apparent connection between ethical facts or properties and motivation. They also help constructivists avoid Street’s skeptical situation. Simply stated, there is no gap between ethical judgment and truth that the skeptic may exploit.²³ But no matter how

²² See Street (2006). She also presents shorter versions of this kind of argument in (2008), (2009), and (2010). Mackie (1977) also presents his own version of the epistemological argument. Street’s is different in that it incorporates evolutionary considerations. It also differs from Mackie’s in that it is intended to support constructivism as an alternative to realism (or quasi-realist expressivism) or an error theory. Mackie’s argument is employed to support error theory, but it has also been appealed to in a different way to support expressivism. Street thinks that she must accept error if constructivism fails. This is because she thinks that her argument cuts equally well against quasi-realist expressivism as it does against realism.

²³ This point needs qualification. As I will explain in Chapter 2, it would appear as if we must distinguish between different orders of stances if we are to make sense of the constructivist’s apparatus. On the one hand, there are the stances that constitute a particular ethical standpoint (what such a *standpoint* is and how we are to understand the nature of the stances that constitute them is controversial; again, see Chapter 2 for discussion). On the other, there are beliefs about what follows from within these ethical standpoints. These beliefs are what gets expressed when we voice an ethical judgment. For example, on this view, we might understand the moral statement “stealing is wrong” – when sincerely uttered by a competent English speaker – to express a belief that certain outcomes follow from within the complex family of stances that constitute her moral standpoint. Different

well one thinks constructivism fares with respect to these naturalistic considerations, this will not be enough to secure it as the preferred alternative to realism. Two problems present themselves.

First, these types of naturalistic concern alone do little to distinguish the constructivist challenge from others, such as the one mounted against realism by error-theorists and expressivists. In fact, it would appear as if every major challenge to realism incorporates some version of this worry.²⁴ So a commitment to naturalism alone would not appear to be enough to force one to accept constructivism as opposed to these other anti-realist alternatives.

Second, the constructivist's specific approach to accommodating naturalism would appear to bring its own costs. Condition (3) implies that there are at least *some* stance-independent ethical standards. Hence, if constructivists reject (3) they are committed to the view that there are *no* stance-independent ethical standards – i.e., that all ethical standards are determined by their ratification from within some actual or hypothetical perspective. In other words, all ethical standards would be *up to us*. Some might worry that this incurs the cost of making the view into an implausible

constructivists present different ways of construing such a standpoint. Some of these will be discussed later. Now, one might worry that this kind of view still allows for skeptical worries to arise because it leaves a gap between an appraiser's beliefs (one kind of stance) and facts about the agent's ethical standpoint (a separate class of stances). In other words, how can we be guaranteed that our ethical beliefs are reliable guides to what is going on amongst our other ethical stances? This picture is complicated further if we consider cases in which the appraiser is not identical to the agent. If – as I will later argue – constructivism is best construed as a version of an agent's group relativism, it turns out that our ethical judgments sometimes concern other people's ethical stances. In these cases, skeptical worries will turn out to be even more pronounced.

²⁴ Here I am putting to one side the question of whether a divine command view in metaethics should be counted as a realist view. Insofar as such a view makes morality, value, or practical reason dependent on God's willing or commanding, one might argue that it too falls into the constructivist camp – or some other camp of non-realist views. However, on most understandings of the *natural* (Moore's excluded), the supernatural is not natural. Hence, we would have a challenge to metaethical realism that would not appeal to realism's putative failure to accommodate naturalism.

form of relativism. Although some recent defenders of constructivism are willing to bite this bullet (in part because they think we can distinguish between objectionable and non-objectionable forms of relativism),²⁵ it is not clear that the naturalistic advantages outweigh this particular cost.

In fact, this kind of tradeoff would appear to be at odds with some of the other motivations that have been offered on behalf of constructivism. For example, early defenders of constructivism argue that their view does a better job of accommodating our deeply-held ethical commitments than realism does. These arguments are often offered as a second type of motivation for rejecting realism.

1.4 Motivations From Within Ethical Theory

In this case, the objection is that realism purportedly fails to accommodate deep features of our ethical thinking. These include the putative connection between moral requirements and reasons for action as well as the importance of autonomy. Unlike the first type of objection, which appeals to our broader philosophical commitments, this one comes from within ethical theory itself.

Some constructivists maintain a *rationalist* view of ethics, according to which, for example, an action's rightness *necessarily* provides some reason to perform it; alternatively, an object's or state-of-affairs' goodness *necessarily* provides some reason to pursue or promote it. These kinds of considerations arguably come from

²⁵ Street (2009) is perhaps most explicit in her acceptance of this potential cost.

within ethical theory, since the relata (i.e., rightness or goodness, on the one hand; reasons for action, on the other) are the proper objects of ethical investigation.

Rationalism per se is not incompatible with ethical realism. In fact, defenders of non-naturalist realist views in ethics – like David Enoch, William Fitzpatrick, Derek Parfit, and Russ Shafer-Landau – also accept something like this view. These philosophers defend a comprehensive metaethical realism (i.e., realism about morality, value, and practical reason). In order to secure an *anti*-realist conclusion, constructivists combine an appeal to rationalism with a rejection of realism about practical reason. This rejection together with a commitment to rationalism puts pressure on one to accept a stance-dependent account of morality and value, as well. Again, their rejection of realism about practical reason in turn either rests on an appeal to broader philosophical commitments, like naturalism (which the just mentioned realists reject),²⁶ or other deeply entrenched moral convictions, like the role of autonomy (which arguably a realist should also be concerned to preserve).

As with the argument from rationalism, the argument from autonomy appeals to a feature from within ethical theory. The claim is that autonomy is an essential feature of *moral*, or *ethical*, agency and that such autonomy requires a kind of control that is at odds with a realist account of ethical objectivity. Insofar as constructivism is designed to accommodate these features, a defender may claim an advantage for her view over realism.

²⁶ Again, some argue on naturalistic grounds that reasons must reduce to something like our motivational states or some other possible objects of scientific investigation.

I devote extended discussion to each of these arguments later in this work.²⁷ But even before we get into the details, one might worry that such arguments will in principle fail to secure their desired conclusions. The problem, in this case, is that no matter how well standard presentations of constructivism fare with respect to these ethical considerations, it is not clear how they show that constructivism enjoys an advantage over *realism*. In short: by appealing to particular, substantive considerations from within ethical theory, constructivists themselves appear committed to condition (3). These ethical starting points are not subjected to scrutiny from within some actual or hypothetical perspective; their status is independent of construction. Hence, it would appear as though these views allow that some ethical truths are stance-independent (i.e., the starting points), while others are not (i.e., what follows from these starting points via construction). If this is correct, it is not clear what the difference between constructivism and realism is supposed to be in the first place. Consequently, it is not clear what there is to be gained by claiming an advantage for constructivism over realism.

One may draw different lessons from this objection. One would be that traditional arguments miss the mark and that constructivists ought to restrict themselves to the naturalistically motivated arguments sketched earlier. The other, by contrast, would be that there are still other ways of drawing the contrast between constructivism and realism that we have not yet explored. In the following section, I will introduce a fourth condition for realism that should help to situate some of the traditional arguments for constructivism.

²⁷ See Chapters 4 and 5, respectively.

1.5 Preserving the Challenge of Substantive Constructivism

It has now become common for both critics and defenders of ethical constructivism to distinguish between substantive and formal versions of the view.²⁸ A *substantive ethical constructivism* is one that characterizes the relevant standpoint for ethical truth in terms of particular ethical claims or judgments (e.g., in terms of what free and equal, or reasonable, persons would agree to, or what would be consistent with valuing humanity in ourselves and others).²⁹ A *formal ethical constructivism*, in contrast, is one that attempts to characterize this standpoint without privileging any particular ethical claims or judgments.³⁰ If formal ethical constructivism is true, then (3) is false. The contrast with realism is straightforward. But the cost, again, is potentially an objectionable form of relativism. I will evaluate formal views against this objection in Chapter 3. The question I will pursue in the remainder of this section is whether there is some condition that distinguishes substantive constructivism from realism.

In the first twenty-five years after its introduction, constructivism in ethics was primarily construed in substantive terms.³¹ For example, Rawls's original position determines what is just for a society by privileging substantive ethical judgments about the value of freedom and equality; Scanlon's contractualist procedure determines what is right and wrong by privileging judgments about what is

²⁸ The terms for this distinction originate in the work of Sharon Street (2008), (2010).

²⁹ On this taxonomy, most of the standard presentations of constructivism count as *substantive* – e.g., those one finds in Rawls (1980), Scanlon (1982), (1998), Korsgaard (1996).

³⁰ Defenders of formal views include Sharon Street (2008), (2010) and Aaron James (2007).

³¹ It is generally accepted that constructivism in ethics first receives expression in John Rawls's *Locke Lectures, Kantian Constructivism in Moral Theory* (1980).

reasonable; Korsgaard's first-person standpoint determines what reasons agents have by privileging the judgment that we have reason to value our own humanity. These views make different classes of ethical truths (those concerning justice, what we owe to each other, reasons for action) dependent on some special class of ethical judgments.

Substantive constructivist views are attractive because of the way they are designed to accommodate central features of our moral thinking. They tap into what are alleged to be some of our deepest-held ethical convictions – e.g., that freedom, equality, and autonomy are valuable; that there is a tight connection between morality and reasons for action; and that these reasons matter.³² Although a realist might deny that these considerations are in fact part of our everyday moral thinking and practice, they have appeared plausible and important to many. From the standpoint of substantive ethical theory, the substantive versions of constructivism appear to have a lot going for them.³³

But, again, it is not clear how these views challenge realism. This is because they allow that there are at least some stance-independent ethical truths. They concede (1), (2), and (3). According to the current consensus, this means that formal

³² Cf. Cuneo (2007). Cuneo proposes that we understand ethical non-naturalism as a theoretical stance that gives primacy to the project of *internal accommodation* (he takes this term from Timmons). Here, I am suggesting that substantive versions of constructivism share this feature. In this light, it is interesting that the major defenders of substantive constructivism (e.g., Rawls and Korsgaard) employ rationalism intuitionism as a foil for the exposition and defense of their own respective views. Rational intuitionism, of course, is one historically prominent example of a non-naturalist moral realism.

³³ However, one may object that substantive views also appear to get things very wrong within in moral theory. For example, these views look as if they will be intensionally inadequate. This is clearest when the constructivist tries to account for certain types of moral judgments – e.g. the judgment that it is wrong to cause gratuitous animal suffering. I will discuss this kind of objection later in the chapter.

constructivism represents the *only* plausible constructivist challenge to realism. I think the current consensus is mistaken.

The objection that a substantive constructivism cannot challenge realism rests on a dubious assumption. It is that only a *thoroughgoing* constructivism can challenge realism – one that involves the rejection of (3) and, with it, all stance-independent ethical truth.³⁴ This assumption can be taken to serve as the first premise in an argument against substantive constructivism.³⁵ The second premise of this argument would be the claim that substantive constructivism is not thoroughgoing; rather, as we have seen, it would appear to involve some combination of stance-independent and stance-dependent ethical truths. From these two premises one must infer that a substantive constructivism cannot challenge realism. But this argument is not sound. This is because the first premise of the argument is false. We should not assume that the only kind of constructivism that can challenge realism is a thoroughgoing one.

The tension between formal constructivism and realism is easy to see. Formal constructivism states that *all* ethical truth is stance-dependent. Realism, by contrast, is compatible with *some* stance-independent ethical truths. Hence, the one position simply commits one to rejecting a constitutive commitment of the other. But the relationship between substantive constructivism and realism is more complicated. This is because both views appear to share certain features. This makes the contrast less

³⁴ *Thoroughgoing Constructivism* is a technical term for Street (2008). My usage is loose and not intended to capture all of the detail that Street assigns to this term. Nevertheless, I take it that what I have to say conforms broadly to her usage.

³⁵ Clarification: The assumption is not that constructivism is the only view that may challenge realism. Presumably the constructivist allows that there are other ways of challenging realism. Rather, the claim is just that if constructivism is to have any chance of undermining realism, it must be thoroughgoing.

apparent. In order to see how these views might conflict, it is first important to acknowledge the features that make them similar.

First of all, realism is perfectly compatible with the view that there are *some* stance-dependent ethical truths to go along with the stance-independent ones. It does not commit one to a thoroughgoing stance-independence.³⁶ For example, a realist will accept that what gives a person an obligation to keep a promise is in some way dependent on that person's being the one who made the promise. This is an obligation that depends on something about the promisor's intentions – e.g., what she has committed herself to do. She would not have an obligation, had she not made a promise.

Furthermore, the realist of course also thinks that there are *some* stance-independent ethical truths and that these truths bear important relations to the stance-dependent ones. For example, in the case of promise making, the truth of a claim about whether an agent has an obligation to keep her promise on a particular occasion will be stance-dependent. It will depend on what exactly she has committed herself to. But, according to the realist, the truth of the general claim that one ought to keep one's promises is arguably not stance-dependent. It is a claim that holds true, if it does, regardless of whether it is entailed from within an agent's ethical standpoint.

Substantive constructivism shares both of these features with realism. It allows that there are both stance-dependent and stance-independent ethical truths *and* that

³⁶ Insofar as realism allows that *some* ethical judgments are stance-dependent in this way, the view also differs from a formal constructivism. Formal constructivism is thoroughgoing. It states that *all* ethical truth is stance-dependence. Realism, by contrast, is not thoroughgoing; it does not commit one to a thoroughgoing stance-independence. Again the assumption underlying the objection to substantive constructivism is that one needs at least one thoroughgoing view to generate a conflict. Since both realism and substantive constructivism are mixed, there can be no conflict between the two. Both allow that there are some stance-dependent ethical truths to go along with the stance-independent ones.

standpoints are structured so that the stance-dependent truths are grounded in the stance-independent ones. For example, on Korsgaard's view, whether we have a reason to do something depends on whether the judgment that we have such reason follows from within the first-person perspective. But the first-person perspective is constituted by certain judgments that are stance-independently true – viz., the judgment that we have *reason* to value our own humanity. These similarities might appear to support the objection that substantive views cannot stand to challenge realism.

But even if we allow that the two views are similar in this way, there still turns out to be an important difference between them. Although a realist may admit of some stance-dependent ethical truths, there is a limit. In other words, realism does not appear compatible with certain stance-dependent truths, regardless of whether they eventually bottom out in stance-independent ones.³⁷

Let us focus on the case of morality since this is arguably the domain in which realist intuitions are most strained by stance-dependent truths. The idea of at least some part of morality being the result of what we will or desire or value is one that is perfectly compatible with moral realism. This much is unproblematic, as the example of promise making illustrates. However, a tension begins to appear when too much of morality is made to depend on the stances one holds. I think this may be illustrated if

³⁷ One might wonder why should we care about this distinction. Even if substantive constructivism is in some sense distinct from realism, it still shares that feature that allegedly makes realism vulnerable to criticism: viz., stance-independent ethical truth. But then won't substantive constructivism be similarly vulnerable to criticism? In short: substantive constructivism might turn out to be distinct from realism in some sense but it cannot challenge realism on the basis of those considerations that constructivists have traditionally brought to bear against realism. This is an important worry, but I think that the substantive constructivist still has grounds to press the challenge. I will explain how this is possible in what follows.

we look at the continuum of substantive constructivisms presented by Rawls, Scanlon, and Korsgaard.

Rawls, Scanlon, and Korsgaard each present different substantive versions of constructivism that occupy varying degrees of the ethical spectrum. Rawls restricts his constructivism to the topic of political and social justice. Scanlon expands constructivism to account for that part of our moral thinking that concerns *what we owe to each other* – i.e., claims about what is right, wrong, or permissible. Although Scanlon’s view also includes a constructivist account of political morality, it allows space for realism about what our reasons are. Korsgaard presents a constructivism about reasons for action. But in virtue of other commitments she makes, her view turns out to take the widest scope, one that expands to include the moral and political, as well.

Korsgaard takes claims about reasons to ground claims about morality (and presumably also political morality). Hence, we may understand her view as assigning a very wide scope to constructivism. However, as I have already argued, it would not appear as if Korsgaard succeeds in capturing *all* of the ethical. For, she must allow that construction bottoms out somewhere. On her view, it is the claim that *we have reason to perform an action if it conforms to principles that would be accepted by a rational agent deliberating from the first-person perspective in ways that are consistent with the demands of agency itself*.³⁸ Contra Korsgaard, this is a claim whose truth would

³⁸ Korsgaard presents a transcendental argument that is supposed to ground this fundamental normative claim without appealing to stance-independent normative reality. However, commentators are divided on this point. Some, like Hussain and Shah (2006), have argued that this kind of bootstrapping does not preclude the truth of traditional metaethical views like realism or expressivism. Street (2010), by contrast, appears to read Korsgaard as arguing for a formal version of constructivism (i.e., one that is

not appear to result from construction. But even if we allow that this one stance-independent truth grounds all other claims about reasons, value, and morality, this should not satisfy a defender of realism – or, specifically, what we might call *robust moral realism*.

Let us take robust moral realism to be the class of views constituted by the satisfaction of (1)-(3) with the addition of one further condition:

- (4) The moral standards that fix our *moral principles* are not made true by virtue of their ratification from within any actual or hypothetical perspective.

A moral principle is any *general* moral statement by which one might judge the moral status of particular actions or states-of-affairs. According to this last condition, the truth of a moral principle is *never* stance-independent. As I will soon explain, this condition would appear to rule out the substantive constructivisms proposed by Korsgaard and Scanlon, while still allowing some particular moral truths to count as stance-dependent (e.g., it allows that one may create obligations to perform particular actions through acts of promise making).³⁹ Hence, it should provide a way of framing the debate between constructivists and a particular kind of realist – viz., a robust moral realist.

Substantive constructivisms fail to satisfy condition (4). Although these views bottom out in stance-independent ethical truths, they allow the content of less basic

thoroughly metaethical and a competitor to realism). Fitzpatrick (2005) appears to accept a similar interpretation but goes on to argue at length that Korsgaard's view fails on its own terms.

³⁹ Whether this condition is compatible with Rawls's *justice as fairness* is considerably more complicated and would require more interpretative work than I am in a position to engage in here.

moral principles to be determined by ratification from within some actual or hypothetical perspective. These “intermediate” stance-dependent moral principles, in turn, provide the criteria by which one might judge the moral status of a particular action or state-of-affairs. For example, on Scanlon’s view, the moral principle *torturing babies for fun is wrong* is true (if true at all) because it would form part of a set of principles for the general regulation of behavior that could not be reasonably rejected by persons complying with the contractualist procedure. Scanlon allows that there are stance-independent truths about what is reasonable. But they do not determine the specific content of less basic moral principles like this one. If they did, the contractualist procedure would prove otiose. The same can be said of Korsgaard’s view. On her view, what makes it the case that this moral principle is true (if true at all) is just that it would be accepted by a rational agent deliberating from the first-person perspective in ways that are consistent with the demands of agency itself. Korsgaard allows that there is a single stance-independent truth about what our reasons are – viz., that we have reason to value our own humanity. But this truth does not determine the specific content of less basic moral principles like this one prohibiting baby torture. If it did, talk of first-person deliberation from within standpoint of humanity would prove otiose. So these paradigmatic versions of substantive constructivism appear incompatible with (4).

However, this condition does allow for *some* stance-dependent moral truths. Again, it was noted that realism is generally taken to allow that, e.g., an agent creates special responsibilities, or obligations, in virtue of her specific commitments. Any condition that ruled out this possibility from the start would be too strong; it would not

pick out a feature of *realism* per se. But (4) does not. Consider, again, the case of promise making. This last condition allows the content of a particular moral statement to be stance-dependent, but it does not allow the content of a general ethical statement to be such. For example, my promising to return you your lawnmower next Thursday (i.e., an intentional act) might make it true that *it is wrong for me not to return your lawnmower to you by next Thursday*. But this statement is not a moral principle; it does not provide any *general* criteria by which we can judge the moral status of any particular action or states-of-affairs other than the one described. Of course, it provides a *particular* criterion. But a particular criterion arguably does not make a principle. So it is not ruled out by condition (4).

Hence, we would appear to have found a condition that allows us to characterize the debate between substantive constructivists and realists while still capturing the ways in which realism is compatible with some moral truths being stance-dependent. But even if this condition does the trick, one might wonder what motivates it. On the one hand, why does realism presuppose the existence of *principles*? Doesn't this condition beg the question against particularist versions of realism? On the other, why think that a *robust* realism about morality requires (4)? What is this notion of robustness supposed to capture? Why should we care? I will quickly say something in response to the first set of questions before turning to the second set of questions in the next section.

One obvious disadvantage of formulating realism in terms of condition (4) is that it rules out *moral particularism* from the start – i.e., views which allow that there are moral truths yet no true moral principles. Condition (4), however, would appear to

presuppose that there are true moral principles, in which case I am begging the question against those who would defend a particularist realism. Although I find this regrettable (since I do not wish to take sides in this debate), it would not appear as if there is a *principled* way of making the distinction between realism and substantive constructivism without appealing to moral principles (no pun intended). The alternative would be to accept that both classes of views allow for some combination of stance-independent and stance-dependent truths but that substantive constructivism allows for *more* stance-dependent truths than realism. This way of drawing the distinction may still be intuitive – in fact, I will soon argue that our intuitions about particular cases provide some independent support for (4) – but its lack of precision is unsatisfying. What’s more, it would not indicate how the dispute might be resolved. *How many more* stance-dependent truths does one have to accept to count as a defender of a substantive constructivism, and not realism? One advantage of the current framework is that it sets a clear line between the two classes of views. Robust realists accept (4); substantive constructivists reject it.

The next set of questions concern whether there is some independent motivation for (4)? If the only reason for thinking that realists are committed to this condition is that it allows them to distinguish their views from substantive constructivism, (4) would be ad hoc. In the next section, I will try to provide some independent reasons for its acceptance.

1.6 Offending Realist Intuitions

Realism is defined by certain theoretical commitments.⁴⁰ But these commitments do not just support themselves. Rather the realist accepts them in part because of the way they appear to support particular ethical judgments in reflective equilibrium.⁴¹ If this is the right way to understand theory building in ethics, we might stand to learn something about the realist's commitments by looking at particular cases.⁴²

Specifically, we might learn something about the differences between realism and substantive constructivism by looking at the kinds of particular ethical judgments that each view supports. Even if both views allow for some combination of stance-dependent and stance-independent truths, we still might take them to give different verdicts on just which particular truths we are to count as stance-dependent.

There is a deep and persistent disagreement amongst philosophical ethicists about which ethical truths are plausibly considered stance-dependent. For example, let us suppose that the following judgment is true: *it is wrong to cause gratuitous animal suffering*.⁴³ I think it is fair to say that this would count as a considered judgment for many, including both realists and constructivists. But people disagree about whether the truth of this judgment is stance-dependent or stance-independent.

⁴⁰ As we have already seen, there is a disagreement about just which commitments realism involves – i.e., whether it involves just (1) and (2), or these and also (3) and (4).

⁴¹ I say *in part* since someone might be inclined to accept certain parts of realism based on her other philosophical or theoretical commitments.

⁴² It is fair to assume that my opponents do think this is right. As I will explain in Chapter 3, every defender of constructivism assigns reflective equilibrium some role in their theory building.

⁴³ Korsgaard (1996) spends a good deal of her final lecture trying to explain how the constructivist can account for norms concerning the treatment of animals. I think this is telling. For it would appear to indicate that she is exercised by this kind of objection.

One could argue that Korsgaard's view, for example, supports the truth of this judgment. But, in this case, the truth of this judgment turns out to be stance-dependent.⁴⁴ Again, according to Korsgaard's view, the truth of this claim and all others, save the one about the reason to value of humanity, are generated by *our* judgments – what follows from within *our* first-person ethical standpoint. Facts about the animal's pain play no direct role in determining ethical truth. At best, the judgments that constitute an agent's ethical standpoint might take facts like these as their objects. But in this case, again, what makes it the case that such actions are wrong is at bottom something about us, not something about the animal or the kinds of experience our actions cause it.

But to some this just does not appear to get things right. These people would argue that what seems to matter most in this case is something *about the animal* – viz., the fact that the animal experiences pain, something that is bad independently of whatever we value, or believe, or desire, etc.⁴⁵ Facts about what we value, or believe, or desire may end up weighing against this concern. But, ultimately, it would appear

⁴⁴ Korsgaard makes this argument in Lecture IV of (1996), esp., 149ff. The argument goes something like this. If this action is wrong then there is some reason for an agent not to perform it. This follows from moral rationalism. If there is some reason for an agent not to perform it, this is because the judgment that one has reason not to is entailed from within the first-person perspective of anyone who values their own humanity. This follows from Korsgaard's substantive constructivist view of reasons. One might question here why a humanity-valuing agent is committed to this judgment. For Korsgaard the answer is supposed to be something to the effect that we share reasons with animals in virtue of our shared animal nature. Any self-respecting human being will respect such reasons for herself and since such reasons are public, she will respect these reasons in the case of animals.

⁴⁵ This point enjoys support from defenders of views from across the metaethical spectrum, including, e.g., expressivists like Simon Blackburn and Allan Gibbard as well as metaethical quietists like Ronald Dworkin.

as if the wrongness of such actions is to be explained by facts that are independent of facts about deliberation.⁴⁶

People will disagree about this case and others like it. But the point here is not that this shows Korsgaard's view to be wrong; rather, it is that this kind of disagreement highlights a fault line that exists between constructivists and realists. Realism may be compatible with the stance-dependence of some ethical truths, but examples like this one illustrate that there is a limit, one that would appear to find expression in condition (4).

This last condition presents a way of explaining and classifying realist intuitions about this case and others like it. It states that there are no moral principles whose truth depends on stance-dependent standards. It follows from this that the moral principle expressed by the statement *it is wrong to cause gratuitous animal suffering* (if true at all) is not made true by stance-dependent standards. In other words, in order to accept (4) one must reject the substantive constructivist accounts advanced by Korsgaard and Scanlon, respectively. Substantive constructivists like Korsgaard and Scanlon are happy to make all moral truths – including true moral principles – stance-dependent. But many moral realists would not be.⁴⁷ My proposal is that this is because these realists accept something like (4). This condition both captures realist intuitions

⁴⁶ In other words, there is still a problem with the intensional adequacy of this kind of substantive constructivism. What's more, one might also ask whether such a view ends up capturing the right extension. For example, Korsgaard's single stance-independent normative truth (that we have reason to value our own humanity and everything that is entailed by this) appears too thin to guarantee the right results.

⁴⁷ Perhaps there are committed realists who will be unmoved by these considerations. If so, this is fine. I restrict my claims to those who wish to distinguish themselves from substantive constructivists. These are people who wish to defend what I am calling *robust moral realism*.

about particular cases and provides a clear target for substantive constructivists, making the substantive constructivist's challenge to realism intelligible.⁴⁸

1.7 The General Argumentative Strategy

As we have seen, constructivists are motivated by two broad classes of considerations – some of these relate to a general commitment to philosophical naturalism, others from commitments within the narrower confines of ethical theory. Defenders of constructivism appeal to these features in different ways. However, as my remarks here illustrate, the arguments they advance conform to a general strategy: first, there is an appeal to a feature from one of the two categories; second, ethical realism is presented as failing to adequately account for this particular feature; finally, constructivism is presented as possessing superior resources for explaining it. In short, the constructivist claims a series of explanatory advantages over realism. Although one might reject one or another of the constructivist's arguments, the constructivist contends that her view wins out on *holistic* grounds.⁴⁹ Even if realism can accommodate some of these considerations to some extent, the constructivist argues that her view does a better job on the whole.

My dissertation starts out from the point of view of a realist who accepts (1)-(4) and is confronted with the constructivist's challenge. On the one hand, the constructivist claims to capture everything the realist truly cares about. As a version of

⁴⁸ If this is right, the next question to ask is what arguments support the rejection of (4)? What kinds of considerations would favor substantive constructivism over realism? These are questions I defer until Chapters 4 and 5.

⁴⁹ See, e.g., Street (2009), 275, 294.

cognitivism, constructivism avoids objections that have traditionally been leveled against views, like expressivism, that deny that ethical claims are the kinds of things that can be literally true or false. As a view that affirms our knowledge of some ethical facts, it avoids the deeply unattractive consequences of an error theory.

On the other hand, constructivism is supposed to get us things that a realist might agree are important but which are difficult to accommodate within a realist theory: It is claimed to provide an account of moral objectivity that is compatible with the best naturalistic theories of reference, epistemology, and ontology. It is supposed to explain the tight connection between morality, or value, and reasons for action. It is also alleged to be the only view that is compatible with our moral autonomy.

On balance, the constructivist claims that her view bears a comparative advantage over realism. Hence, we have the constructivist's challenge to realism: if constructivism accommodates everything good about realism *and more* while avoiding standard objections, what more could the realist's extensive commitment to stance-independent ethical facts have to add to the picture, except additional problems? It is its proximity to realism that makes constructivism a serious challenger.

I count myself as someone who is sympathetic to commitments (1)-(4) but also to the kinds of considerations that constructivists marshal in their challenge to realism. In order to stand firm in these commitments, it is important to see how a realist might respond. It is with this motivation that I embark on a defense of ethical realism.

As constructivist arguments make claims against realism and on behalf of constructivism, my defense of realism also involves both positive and negative arguments. On the one hand, I argue that realism is able to accommodate some of the

appealed to considerations better than constructivists claim it can; on the other, that constructivism does worse than advertised. In the end, my approach is also holistic. But my conclusion is that constructivists provide no good reason for thinking their views bear an advantage on balance over ethical realism.

One may dispute whether the commitments enumerated by the constructivist match our actual thinking and practice. This would be one way of meeting the constructivist's challenge. However, this approach represents the road not taken here. Instead, I proceed by first granting the constructivist these starting points and then arguing that the constructivist's conclusions do not follow from them.

The discussion of realism and constructivism in this introduction provides a framework for the chapters to come. Conditions (3) and (4) have to do with the nature and extent of ethical objectivity. In Chapter 3, I will evaluate the position maintained by those formal constructivists who reject (3). In Chapters 4 and 5, I turn to some of the traditional arguments that have been advanced by substantive constructivists who reject (4). These three chapters contain my evaluation of the constructivist's challenge. However, before we begin with this, it is important that we address a few remaining interpretative issues. In Chapter 2, I will examine the question whether either version of constructivism – formal or constructivism – should be considered a free-standing metaethical alternative to already familiar positions.

CHAPTER 2

IS CONSTRUCTIVISM A *FREE-STANDING* METAETHICAL VIEW?

2.1 Introduction

For all that I have said thus far, one might still wonder what is supposed to make the constructivist's challenge a special one. There are a couple of worries here. First, regardless of which view (realism or constructivism) provides the better account of these various considerations (naturalistic and ethical), mightn't there be other views out there that do just as well as (or better than) either (or both) constructivism or (and) realism? This is surely an important question. But its scope is beyond what I could hope to achieve in this work. As I already started to explain in the last chapter and will further specify in the current one, constructivism poses an interesting and important challenge to realism that deserves its own attention. If I can show that realism can stand up to this challenge, this will contribute something important to our understanding of metaethical debates, the comparative success or failure of other views notwithstanding.

Second, one might worry that 'constructivism' is merely a new label for another well-established view, one whose virtues and defects have already received much attention. Specifically, much of what I have said thus far about constructivism (*both* formal and substantive) appears as if it applies to a family of views that are sometimes referred to as *response-dependence theories*. More recently, both constructivists and non-constructivists alike have questioned constructivism's

cognitivist credentials and pressed for details as to how such a view might contrast with expressivism. Furthermore, one might worry that constructivists only succeed at distinguishing their views from expressivism (or response-dependence theories) to the extent that they construe them as a form of simple subjectivism such a *desire-satisfaction theory*, i.e. a naturalistically reductive view that make ethical facts a function of an agent's desires. How is ethical constructivism different from such views? Is constructivism a distinct, or *free-standing*, alternative to response-dependence theories, expressivism, or simple subjectivism? Or does it perhaps represent a species of one of them? Alternatively, might we take constructivism to be the family or genus under which these other views fall?

If constructivism merely turned out to be a version of one of these other views, this would arguably detract from its importance. Part of what is supposed to make the constructivist challenge to realism an interesting one is that it has not received the attention it deserves. This would arguably not be true if it turned out to be a version of a response-dependence theory, or expressivism, or simple subjectivism. Although the commitments of these other views are still controversial, they have each commanded a lot of discussion already. In light of this, the strengths and vulnerabilities of each type of view are fairly well established. What's more, it is difficult to see how constructivism might serve as an improvement on any of these views – since the objections to each, respectively, are general enough that they would appear to extend to any of their species, constructivist or otherwise. For this reason, it is important to ask whether there is a sense in which constructivism might count as a *free-standing* metaethical alternative to these views.

In this chapter, I set out to address this question and show just how difficult it is to fashion a metaethical version of constructivism that remains distinct from familiar positions. In sections 2 and 3, I review standard proceduralist characterizations of constructivism and argue that they are best understood as just another version of a response-dependence theory. In sections 4 and 5, I present recent attempts to characterize moral truth as the outcome of non-procedural standpoints and show that these fold into realism, expressivism, or a desire-satisfaction view – none of which would mark a new position. Furthermore, I argue, there is no reason to think that constructivism improves on these other views in ways that make for a more formidable challenge to realism. In short: constructivism does not turn out to represent a free-standing metaethical alternative to realism. This conclusion prompts me to consider other ways in which constructivism might challenge ethical realism. I consider these alternatives in section 6.

2.2 The Proceduralist Characterization

Constructivism has typically been understood as the view that ethical truth is determined by the outcomes of certain “suitably described” procedures. A view that characterizes constructivism in terms of procedures easily lends itself to formulation in terms of a response-dependence theory. As such, constructivism would appear to represent a species of such views and, consequently, be subject to the same

objections.⁵⁰ The *proceduralist characterization* is in large part due to the influence of John Rawls.⁵¹ In the course of developing a foundation for debates about political justice that could be divorced from what he considered deep and irreconcilable metaphysical disagreements, Rawls (1980) introduced a conception of constructivism in moral theory.⁵²

[Moral] constructivism holds that moral objectivity is to be understood in terms of a suitably constructed social point of view that all can accept. Apart from the procedure of constructing the principles of justice, there are no moral facts. Whether certain facts are to be recognized as reasons of right and justice, or how much they are to count, can be ascertained only from within the constructivist procedure, that is, from the undertakings of rational agents of construction when suitably represented as free and equal moral persons.⁵³

On this view, ethical truth is *determined* by procedures in the sense that the ethical standards that fix the relevant class of ethical facts are constituted by their emergence

⁵⁰ Both David Enoch (2009) and Michael Ridge (forthcoming) argue, for this reason, that constructivism cannot be a free-standing position. Both James (2007) and Street (2010) recognize the problems with the procedural characterization and work to establish a formulation of constructivism that avoids this. I discuss Street's alternative characterization shortly.

⁵¹ I take this label and that of its alternative – *the standpoint characterization* – from Street (2010), 364.

⁵² Rawls's first presentation of constructivism in (1980) is arguably to be understood as a comprehensive ethical outlook – one that describes the nature of ultimate ethical reality (including the political). As such, it does not represent Rawls's considered view of political justice. In contrast, Rawls's later work is characterized by his defense of a position he called *political constructivism* (see, e.g., Rawls (1996)). This view, in contrast with his earlier constructivism, is explicitly intended to be neutral with respect to one's deeper religious and metaethical commitments (i.e., what Rawls calls *comprehensive doctrines*). According to Rawls, it is this neutrality that is central to any view that is to have a chance of providing a foundation for stability in a political society of reasonable citizens who yet fundamentally disagree on issues of religion and metaphysics. For a discussion of Rawls's development see Samuel Freeman's "Introduction" to Rawls (1999).

⁵³ Rawls (1980), 519.

from special procedures. It is these facts that make our ethical assertions true or false. It is not the case that these standards merely coincide with such outcomes or that the procedures may be used to discover what the relevant standards are. For example, Rawls famously argues that facts about justice are fixed by principles that would be agreed to in the *original position*, a procedure in which free and equal citizens agree to terms of social cooperation under the condition that they do not know facts about who they are or what they deeply care about.⁵⁴ Apart from such a procedure, Rawls claims that there are no facts about justice.

The proceduralist characterization of constructivism has been accepted both by advocates, like Milo (1995) and Korsgaard (1996), and by critics, like Darwall, Gibbard, Railton (1992), Enoch (2009), and Ridge (forthcoming).⁵⁵ Again, this has led many to doubt whether constructivism represents a free-standing metaethical position. In particular, it is natural to think that constructivism's reliance on procedures makes it a version of a response dependence view in metaethics.

⁵⁴ This condition on choice is what Rawls calls *the veil of ignorance*. See Rawls (1971).

⁵⁵ See, e.g., Milo (1995), 186. Korsgaard (1996) even refers to constructivism as *procedural moral realism*:

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

2.3 Constructivism as a Response Dependence View

Response-dependent views have been described in different ways, and many debates about them are still open.⁵⁶ Some have defended a response-dependence view of ethical properties. We might take such views to provide the following schematic account of the *essence* of ethical properties:

x is C iff (and because) x is such as to produce R in S s under conditions K

where C stands in for some ethical property (e.g., the property of *goodness*, *wrongness*, *being a reason*), S the subject, K the relevant conditions, and R the response.⁵⁷ In order for this equation to pick out a response-dependent property, there are other conditions that must also obtain.⁵⁸ But, sidestepping the controversies involved in specifying such conditions, it is already apparent why one might take constructivism to represent a species of such views. The proceduralist characterization lends itself to formulation in terms of subjects, responses, and conditions. For example, Rawls's statement of a constructivist view of justice might be made to fit the response-dependence schema in the following way.

⁵⁶ See, e.g., Mark Johnston (1989), (1993), David Lewis (1989), Michael Smith (1989), Crispin Wright (1988), and Philip Pettit (1991).

⁵⁷ This formulation of response-dependence is loosely based on the view presented in Johnston (1989). There are a couple of differences. Perhaps the main one is that Johnston's view is a theory of *concepts* not properties. It is unclear what such a view of concepts commits one to metaphysically. As I am ultimately interested in constructivism, a view about the metaphysics of ethical facts and properties, I have gone ahead and re-formulated Johnston's schema to provide a response-dependent account of ethical properties.

⁵⁸ For example, the bi-conditional cannot obtain trivially in virtue of S , K , and R specifying a "whatever it takes" condition.

A policy/institution, x , is *just* iff (and because) x conforms to principles that would be agreed to by free and equal citizens under the conditions represented in the original position.

Here, C is the property of being just. R is the disposition to produce agreement, a volitional response. S is a free and equal citizen of a society. K is the conditions described in the original position. There may be renderings of Kantian constructivism about justice that better capture Rawls's view.⁵⁹ My point is just to illustrate how a view that focuses on procedures lends itself to an interpretation in response-dependence terms. In fact, Rawls's statement of constructivism is not unique. The constructivist proposals of T.M. Scanlon and Christine Korsgaard are also amenable to a response-dependence formulation. For example, Scanlon's *moral* constructivism ("what we owe each other"), might be captured thusly:

Φ -ing is *wrong* iff (and because) Φ -ing is disallowed by any set of principles for the general regulation of behavior that could not be reasonably rejected by persons complying with the contractualist procedure.

Here, C is the property of wrongness. R is being unable to be reject a set of principles for the regulation of behavior, again (arguably) a volitional response. S is a person

⁵⁹ One may argue, e.g., that Rawls intends the conditions of freedom and equality to be "modeled" by the original position, in which case they should be taken to specify K – not S ; or that S should include a description of the citizen's motivations – i.e., that they are motivated to find terms of agreement that other, similarly motivated citizens, can reasonably be expected to uphold in good faith.

motivated to comply with outcomes of the contractualist procedure. *K* includes the conditions that make up such a procedure – i.e., that the agreement reached be informed and unforced and that others similarly motivated could also not reasonably reject it.⁶⁰ Similarly, Christine Korsgaard’s constructivism *about practical reason* might take the following response-dependent formulation:

A feature of an action, or a situation, *x*, counts as a *reason for acting* iff (and because) *x* conforms to a principle that would be accepted by a rational agent deliberating from the first-person perspective in ways that are consistent with the demands of agency itself.

In this case, *C* is being a reason to act. *R* is the disposition to accept a principle. *S* is a rational agent. *K* is the conditions required to maintain agency within the first-person perspective.

But if ethical constructivism is best understood as a response-dependence view, as these examples suggest, it would not represent a free-standing metaethical view. What’s more, it would appear to be subject to the same objections that have been leveled against these views. Different philosophers have recognized a dilemma for response-dependence theories.⁶¹ It is this. We may understand a response-dependence account as providing either a non-reductive or a reductive account of

⁶⁰ Scanlon (1998), 153-4.

⁶¹ See, e.g., Blackburn (1993b), (1998), Cuneo (2001), Wright (1988), and Zangwill (2003). As I mentioned earlier, Darwall, Gibbard, Railton (1992) understand constructivism to be a kind of hypothetical proceduralism. Consequently, they argue that the view is subject to this kind of dilemma (see esp., p. 143). For a collection of articles on response-dependence views, see Casati and Tappolet (1998).

some class of ethical properties. If it is *non-reductive* (i.e., it includes some ethical terms on the right-hand side of the equation), the account will leave traditional metaethical questions unresolved. Specifically, it will not tell us how we are to understand the ethical terms employed in the account, leaving open the possibility that they may take a realist, or expressivist, or error-theoretic, etc., interpretation. If it is *reductive* (i.e., it includes no ethical terms of the right-hand side), the account runs the risk of being extensionally inadequate. In other words, the outcomes of the reductively specified procedure will not match our considered ethical convictions. In the case of constructivism, this might be expressed as the worry that the subjects of the relevant procedures, actual or hypothetical, could get things wrong.⁶² In fact, this particular worry may also apply to non-reductive accounts as well, especially if one is careful to avoid “whatever it takes conditions” in specifying the procedure.⁶³

A defender of the proceduralist characterization might argue that constructivism provides some new way of navigating around this dilemma. But it is difficult to see how it could. The features of a response-dependence view that make it vulnerable to the dilemma are general and appear, as such, to apply equally to a proceduralist interpretation of constructivism. A more promising response, perhaps,

⁶² What’s more, reductive accounts are also subject to the worry that they are *intensionally* inadequate. In other words, they do not accurately express what ethical thought and practice is intuitively *about*. However, this kind of objection generalizes to just about any substantive account of the nature of ethical properties and facts. It is for this reason that I will focus on the extensional worry. Here the worry is not just that the account commits one to outcomes that might diverge from common sense (again, such a worry might apply to any reductive account); rather, it is that the way in which response-dependence theories route correct outcomes through a subject’s intentions opens these views up to error in a special way.

⁶³ For example, if *right* actions are those that conform to standards decided on by paradigmatically *good* agents (e.g. Pericles, Jesus, Gandhi, King, etc.), we might still ask whether such agents could get things wrong. And, in fact, to the extent that these figures are *mere* men, it would appear as though some error is inevitable.

would be for the constructivist to accept one of the horns but argue that the associated objection is not as bad as one might think. But this type of response is also available to defenders of response-dependence views more generally. It still would not provide us with any reason for thinking that constructivism was a free-standing metaethical view; rather, it would appear as if constructivism and response-dependence theories (in general) stand or fall together. If constructivism is to represent a free-standing view it cannot be construed in terms of procedures. This has led more recent defenders of the view to reject the proceduralist characterization and emphasize the ways in which constructivism differs from response-dependence theories.

2.4 The Standpoint Characterization

Sharon Street (2010) takes constructivism to represent a distinct and superior alternative to response-dependence views in metaethics. On her most recent account, she describes metaethical constructivism as the view that

...the truth of a normative claim consists in that claim's being entailed from within the practical point of view, where the practical point of view is given a *formal* characterization.⁶⁴

⁶⁴ Street (2010), 369. NB Street appears to use the term 'normative' in different ways. Sometimes she uses it to refer to a general category that captures morality, values, and reasons (what I am calling *the ethical*). Other times she appears to use it in a restricted sense to refer to reasons for action alone.

Street's account explicitly avoids any kind of procedural specification. Again, if constructivism specifies a procedure, this leaves the view open to the dilemma just sketched. From this one might infer that the solution is to avoid specification.⁶⁵ So, the constructivist retreats from talk of procedural outcomes to what is alleged to be the less problematic, talk of *points of view*, or *standpoints*.⁶⁶ According to this kind of view, when we are evaluating the truth of an ethical judgment, we are not concerned with any particular subject's responses (hypothetical or actual); rather, we are concerned with what follows from within a particular framework for engaging the world. This framework is not an agent. Consequently, the worry that a subject's responses might get things wrong is not supposed to not apply.

But what is a standpoint? Although many philosophers appeal to standpoints (especially, those working in the Kantian tradition), there is very little detailed discussion of what they are. For example, Street describes the *practical point of view* as

...the point of view occupied by any creature who takes at least some things in the world to be good or bad, better or worse, required or optional, worthy or

⁶⁵ James (2007) makes the same adjustment in light of this objection. The main difference between the formal view he proposes and the one described by Street concerns whether one can expect a convergence between various agent's standpoints. This difference is discussed in more detail below.

⁶⁶ Street illustrates the difference by appeal to an example about baseball and how response-dependence and constructivist views would differ in their response to the question *What is it for Jeter to be safe?* Whereas the former would state the conditions for Jeter's being safe in terms of the responses of an umpire (e.g., a good umpire would judge him to be safe), the latter states the conditions in terms of what would be entailed by the rules of baseball in combination with the non-normative facts. Street argues that this formally-construed constructivist alternative is immune to the standard objections that befall response-dependence accounts. On the one hand, it yields the right results; it is extensionally adequate. Unlike the response-dependence view it leaves no room for errors of judgment. On the other, it would appear to capture the sense of *what it is* for Jeter to be safe; it is also, on Street's view, intensionally adequate.

worthless, and so on – the standpoint of a being who judges, whether at a reflective or unreflective level that some things call for, demand, or provide reasons for others.⁶⁷

This description presupposes that we already start out with some sense of what such a standpoint is. It is not informative. Other descriptions of standpoints appeal to metaphor or invoke a distinction between the practical and the theoretical, each of which is supposed to represent a distinct and familiar way of experiencing the world.⁶⁸ The idea appears to be that we all occupy these standpoints and have some intuitive sense of what they are.⁶⁹ But surely there is more to say. In what follows, I sketch out in my own way what I take a standpoint to be and how the constructivist might employ it to account for ethical truth. This account is based on the ones given by Street, James, and Korsgaard. However, as each of these constructivists describe things in different ways, emphasizing certain aspects and neglecting others, I find it necessary to present the view in my own terms, smoothing out the contours and filling in the gaps that characterize these other accounts.

⁶⁷ Street (2010), 366. This description is consistent with the way that she has described things in earlier work. For example, compare this with her description of a practical standpoint in (2008):

A practical standpoint is the standpoint of one who makes normative judgments. One occupies the practical standpoint whenever one judges that some things provide practical reasons, or are valuable, good, bad, required, worthwhile, and so on. (p. 209n)

⁶⁸ For this approach see, e.g. Wallace (2008).

⁶⁹ The operative principle here is something like “you’ll know it when you see it.” One reason that these accounts may be lacking in detail is that a standpoint is presumably a primitive structuring of experience and, as such, hard to get “outside of” and describe. The point of the metaphors is to help us attend to the features that make up a particular perspective “from the inside”. But, even if a standpoint is a primitive structure that informs our most basic experience of the world, there must be more that we can say about it. If it is indeed *structured*, we should at least be able to say something about what this structure is like.

Let us assume that a *standpoint* is constituted by a complex system of intentional psychological states (i.e., stances), such as beliefs, desires, commitments, reactive attitudes, etc. In other words, it is a set of individual stances that hang together in a certain way. This much would appear safe if we are correct in understanding the distinction between realism and constructivism in terms of stance-dependence. But if we are to avoid the response-dependence dilemma, it is important that these stances not be described as issuing from any particular type of subject under specific conditions. Again, the worry about extensional inadequacy would appear related to the idea that our subject could get things wrong. Alternatively, one might focus on the kinds of activities associated with various practical standpoints. Here the idea is that we first look to those familiar activities of, e.g., *valuing*, *taking something to be wrong*, *taking something to be a reason for acting*, in order to identify the relevant practical standpoint and then ask what it is to engage in them *as such*.

Korsgaard (2003), James (2007), and Street (2010) describe the constructivist project as one of working out the *constitutive commitments* of various practical standpoints. This involves the task of specifying the various ways in which particular types of stances must hang together so that one may count as genuinely engaging in a particular practical activity. For example, stances can presumably either support or conflict with one another. Conflicting stances are ones that are in some sense *inconsistent* with each other.⁷⁰ Although Katie may consistently take herself to have both some reason to attend the concert and some reason not to, one might think that

⁷⁰ The sense depends on how one is to understand the nature of the stance's content. As I will soon explain, some have argued that once the nature of this content is specified constructivism reveals itself to be merely a species of already well-known metaethical views – either realism or expressivism.

she may not consistently take herself to have both an all-things-considered reason to attend and an all-things-considered reason not to.⁷¹ Someone who simultaneously maintains these stances arguably fails to meet the basic requirements for *taking something to be an all-things-considered reason for acting*. This kind of conflict illustrates perhaps the most straightforward sense in which practical stances may count as inconsistent. But there are other ways, too. Consider someone who takes herself to have an all-things-considered reason to save her own life, believes that in order to do so it is necessary that she see a doctor immediately, but takes herself to have *no* reason whatsoever to see a doctor. These stances are also in some sense inconsistent with one another. Like in the earlier example, someone who simultaneously maintains these stances arguably fails to meet the basic requirements for *taking something to be an all-things-considered reason for acting*. A standpoint constructivist might claim that this is because the activity *as such* is in part constituted by a norm of instrumental consistency, one that requires that one take oneself to have at least some reason to take the necessary means to one's ends. As both Street and Korsgaard are wont to point out, someone who fails in these ways is not making a mistake about what her reasons are; rather, she does not genuinely count as taking herself to have an all-things-considered reason at all.⁷² But consistency is not the only kind of relation that one might take to matter.

⁷¹ Some may wish to contest this since it would appear to preclude the possibility of genuine dilemmas in practical reason. But this need not matter here. I only intend to provide an example of the kinds of things that constructivists could mean when they talk about a formal constraint of consistency. What's more, this particular kind of example is one that Street, James, and Korsgaard all appeal to.

⁷² Street (2008), 228-229; (2010), 374. Korsgaard defends this claim across a number of works, including, e.g., (1997) and (2009).

Those stances that do not conflict may stand in various degrees of support to each other. Amongst a set of mutually consistent stances, some will be more central to a particular standpoint, others more peripheral. The extent to which stances, on balance, exhibit support of one another is a measure of their *coherence*. By comparison with consistency, it is more difficult to say how coherence might figure as a constitutive requirement for a particular standpoint. Presumably, one counts as occupying a standpoint or one doesn't. It is an all or nothing affair. But coherence comes in degrees. Surely one may count as genuinely engaging various practical standpoints even if the relations amongst one's set of stances falls short of maximal coherence. But what if they fell short of *minimal* coherence?

Coherence might figure as a threshold requirement. Consider someone who is deliberating about whether to attend a party. She considers who will be there, whether there will be dancing, how she will feel the next morning, how this might affect her work schedule, etc... After much reflection she concludes "I have an all-things-considered reason to attend the APA this year!" Although there is nothing apparently inconsistent with her taking up this stance, it does not mesh in any way (let us assume) with the considerations she has been entertaining. It is wildly incoherent with them. Someone who takes up this stance in the present situation arguably fails to count as taking herself to have an all-things-considered reason. Here we might say that the relation this stance bears to the background of the agent's other stances exhibits a degree of coherence that falls below the threshold that is constitutively required to be counted as engaging in the activity as such.

Coherence might not be the only relation that matters in this way.

Alternatively, there may be different constitutive norms whose satisfaction contributes to a standpoint's overall coherence. Part of the constructivist project will involve describing what other types of norms or relations are constitutive of various practical standpoints. For example, James (2007) has claimed that standpoints of practical reason are constituted by certain general norms which govern, amongst other things, which facts one is to attend to in deliberation (a "norm of attention direction"), which to disregard (a "norm of disregard"), which to count as favoring a particular response (a "norm of favoring"), and which to assign more or less importance (a "norm of balancing").⁷³ Like the norm of coherence, but unlike the norm of instrumental consistency, these norms would also appear to allow satisfaction to various degrees.⁷⁴

Once the constructivist has an account of these various constitutive commitments in hand, she can then appeal to them to explain the truth, or falsity, of a particular ethical judgment. According to the standpoint characterization of constructivism, the truth of an ethical judgment is a function of what *follows* from

⁷³ James (2007), 315-19. For James it would be appropriate to speak of *the* standpoint of practical reason because he thinks that there is a necessary convergence in ethical truth, regardless of the stances one starts out with. One might also say the same of Korsgaard's view - though her view counts as a substantive constructivism, by contrast with James's formal view. However, for Humean constructivist's like Street it is more appropriate to speak of *a* particular standpoint since, on her view, the entailments from a standpoint are largely determined by an agent's starting points, and these (she allows) may differ significantly across agents. In fact, Street is explicit about the relativistic implications of her view. See, e.g., Street (2008), 224. But she argues that whatever relativism is involved is not objectionable. For an extended argument, see Street (2009).

⁷⁴ For example, one may succeed, or fail, to attend to the right features of a situation to greater or lesser extents. Someone who is trying to decide whether or not knowledge has intrinsic value might be said to attend to the wrong kinds of facts if she devotes all of her attention to facts about the practice of astrology or wizardry. But she arguably does better if she does this than if she attends solely to the perceived redness of an apple. Astrology and wizardry might be pseudo-sciences but they bear a closer relation to the topic of knowledge than the redness of an apple. Cf. Street (2008), 235. There Street claims that the truth of a particular normative judgment "...is a function of how strongly [an agent] holds the normative judgment in question and how close to the center of her total web of normative judgments they lie" (p. 235).

within a particular practical standpoint. Given a particular collection of stances and the norms governing a particular standpoint, certain stances will follow, others will be ruled out, and yet others will fall somewhere in between, enjoying some degree of support but falling short of “entailment”.

We might ask whether constructivism, on this view, counts as a kind of ethical relativism – one that relativizes ethical truth to the features of the standpoint of the agent performing the action, or confronted with a particular state-of-affairs. The answer will depend on how one characterizes the standpoint as well as the view we are calling *relativism*. For example, there are some defenders of this view who also think that the constitutive features of a standpoint guarantee convergence on a single ethical class of ethical truths. This version would appear to speak against the way in which some have interpreted relativism.⁷⁵ The charge of relativism notwithstanding, this sketch allows us to see how constructivists, like Street, might appeal to a standpoint

⁷⁵ Cf. Harman (2001). This brief sketch leaves this and many other questions unanswered. Here are just a few other questions: (1) Are all three ethical standpoints commensurable or do they express different states of mind? If they are necessarily connected – as *rationalism* claims they are – it would seem that they must express the same kind of state of mind. The basic story will be about the attitude of *taking something to be a reason*. Accounts of the other ethical attitudes will be built up out of this. If you reject rationalism, as Street does (see Chapter Two), there is no pressure to account for them to express the same kind of state. (2) Will the constitutive commitments of one kind of attitude provide enough guidance for determining the correctness of more than a handful of judgments? Won't there be indeterminacy all over the place at a certain stage of development? Suppose that I start out with a quorum of stances, i.e., whatever number suffices to constitute a standpoint. As I acquire new stances it would seem that the constitution of the standpoint could go in any number of directions. Suppose I am entertaining the stance *P*. Perhaps *P* cleaves my current set of stances in a way that previously was not apparent – because I hadn't thought about the kinds of issues that *P* relates to. Suppose that half of my current stances support *P*; the other half conflicts. Do we accept *P* or not *P* as a stance? Whichever one is chosen will end up affecting the constitutive entailments of the standpoint going forward. But the worry is that the choice is indeterminate. What's more, the objection is not merely that there will be indeterminacy at some points. It is that we are faced with these kinds of choice points from the very beginning. How *ought* we to go about acquiring new stances if this is the case? (3) Must one assume that a standpoint will be constituted by just one type of stance? Couldn't, e.g., the moral standpoint be constituted by a complex system of beliefs, desires and other attitudes?

characterization to distinguish their views from the family of response-dependence theories.

Our judgments about a particular practical standpoint, importantly, are not *about* how some subject – actual or hypothetical – would respond. The dilemma for response-dependence views is a result of the way in which these views describe ethical truths *indirectly*. On these views, the truth of an ethical judgment is described in terms of how some suitably described agent would respond. This allows for a gap between our intuitive ethical judgments and our judgments about how such agents would respond. On the one hand, the agent could get things wrong; on the other, our ethical judgments themselves (the direct ones) might not appear to be *about* such agents and their judgments. The move to standpoints is supposed to close this gap.

According to the standpoint characterization of constructivism, an ethical judgment isn't about what some suitably described agent would judge or choose.⁷⁶ Rather, it is about what follows from within a particular ethical standpoint. In contrast with response-dependence views, there is no intuitive gap between what follows from within the standpoint and ethical truth. The standpoint does not represent an agent. Hence, there is no one who could get things wrong. What's more, such a standpoint

⁷⁶ Cf. Street (2008):

We are not asking what *A* or anyone else *thinks* withstands scrutiny from that standpoint. Rather, we are asking whether, as determined by the standards set by *A*'s other normative judgments in combination with the non-normative facts, the judgment that *X* is a reason to *Y* (for *A*) *does* withstand scrutiny from that standpoint. (p. 231)

does not reduce ethical judgments to anything else; this is supposed to make it immune “open question” worries about the accounts intensional adequacy.⁷⁷

2.5 Problems for the Standpoint Characterization

One might worry that the extent to which the standpoint characterization succeeds in distinguishing constructivism from response-dependence views is also the extent to which the view starts to look like other existing metaethical alternatives, viz. realism, expressivism, or a simple subjectivism. Suppose that what distinguishes constructivism from response-dependence views is that on the former view ethical judgments are *about* what follows from within a particular standpoint and not how a certain type of subject would respond under suitable conditions. For example, compare how the two views would characterize the moral statement *stealing is wrong*.

According to a response-dependence view, the statement means something about a particular type of subject’s responses; according to the standpoint view, the statement means something about what follows from within a standpoint constituted by certain attitudes and formal constraints.⁷⁸ The standpoint view makes it is clear how constructivism might still count as a version of cognitivism. Ethical judgments are a species of belief, ones that report facts about what follows from within a particular ethical standpoint. But how are we to understand the nature of the relevant stances that

⁷⁷ The standpoint account leaves many questions unanswered about the nature of ethical facts and properties. Is the standpoint account supposed to support our everyday moral practice, or undermine it and prompt revision? What kind of epistemology is required on such an account? Can such an account be accommodated within a naturalistic framework? These are the kinds of questions that one might have thought constructivism was supposed to provide answers to.

⁷⁸ Again, different constructivists characterize the role of initial attitudes as well as the kinds of constitutive conditions that make up a particular kind of standpoint in different ways.

make up these standpoints? So far, all that we have been told is that they are not about a subject's responses.

According to the standpoint view, we might say that ethical judgments involve different orders of stances. On the one hand, when a competent English speaker utters an ethical statement (like *stealing is wrong*), she is expressing a belief. This belief is about what follows from within a particular agent's ethical standpoint. There are two orders of stances. On the one hand, there are the stances that make up the relevant standpoint. These belong to the agent. On the other, you have the appraiser's belief (a type of stance) about this standpoint. This means that we must distinguish between two types of cases. In the first kind, the appraiser is herself the agent, in which case her beliefs are about her own ethical standpoint. In the second kind, the appraiser is not the agent, in which case the appraiser's ethical judgment expresses a belief about someone else's standpoint. Importantly, on this kind of view, ethical truth is always relative to a standpoint. However, whether there will be divergence in ethical truth across agents depends on how one characterizes such standpoints.⁷⁹

On this view, it is clear that the appraiser's ethical judgments express beliefs *about* what follows from within the agent's ethical standpoint. But how are we to understand the agent's stances – the ones that make up a standpoint? What are they? Are they belief-like states with their own representational content? Or are they non-belief-like states? These questions are crucial for determining whether constructivism represents a free-standing view. However it is not clear that the standpoint

⁷⁹ I'll say a bit more about this later in this discussion. For a fuller discussion, see Chapter 3.

constructivist can answer them in a way that succeeds in distinguishing her view from familiar alternatives.⁸⁰

Suppose that the stances that make up an agent's ethical standpoint also have representational content and a belief-like *direction of fit*. That is, we aim to conform these attitudes to information we receive about the world.⁸¹ We might call these *standpoint beliefs* to help us to distinguish them from the beliefs expressed by the appraiser in her ethical judgments. What is the content of such beliefs? Alternatively, what would have to be the case for these standpoint beliefs to be true?

If they are made true by features of the world that obtain independently of our stances, the view just turns out to be version of realism. For, in this case, there are at least some ethical judgments whose truth does not ultimately depend on the relations they bear to other stances within a particular ethical standpoint; rather, some will ultimately depend on the relation they bear to the world. For all I have said, the constructivist apparatus may still be able to contribute something novel and interesting within such a framework. For example, even if constructivism turns out to be compatible with realism in some ways it may serve to undermine it in others.⁸² I devote extended discussion to this possibility in Chapter Three. On the other hand, one may argue that this objection overlooks the possibility that the content of a first-tier belief involves essential reference to other beliefs that constitute the standpoint. Such a

⁸⁰ My thinking on these issues has been influenced and (to my mind) benefited from some, as yet unpublished, work that Michael Ridge has done on this topic. See Ridge (forthcoming).

⁸¹ Anscombe (1957)

⁸² This is a possibility that I sketched out in Chapter 1 when discussing substantive constructivism. See Chapters 4 and 5 for an evaluation of arguments to this effect.

view may avoid realism. But, in this case, the cost is that this makes the view either vacuous or objectionably circular.⁸³

According to this second proposal, we are to take every basic belief that constitutes a standpoint to have content that makes essential reference to other beliefs constituting the standpoint.⁸⁴ This means that no standpoint belief will be true simply in virtue of stance-independent features of the world. It also means that the content of many of our everyday ethical beliefs will turn out to be a bit different from what they might appear. For example, on such a view, we never merely have the belief that *stealing is wrong*; rather, ethical beliefs are always relative to a standpoint. Explicitly or implicitly, the content of an ethical belief must make essential reference to the other beliefs that constitute the relevant agent's moral standpoint. For example, it would have to be something like the belief that *the claim that stealing is wrong enjoys support from the beliefs that make up the agent's moral standpoint*. One problem here is that it is hard to see how this says anything about morality. On this view, someone who sincerely utters the statement "Stealing is wrong" is expressing her belief that the belief that the claim that stealing is wrong enjoys support from the beliefs that make up the agent's moral standpoint. This account appears vacuous. It says that basic ethical beliefs are about other beliefs, but it doesn't say anything more about what the content of those beliefs are. What would it be for the other beliefs that make up the agent's moral standpoint to support the claim that stealing is wrong? Presumably, these other beliefs will in turn reference the support of other beliefs that constitute a

⁸³ See Ridge (forthcoming) for the circularity objection. James (2007) also recognizes and responds to something like this objection.

⁸⁴ Cf., Brink (1989) on Rawls's presentation of Kantian constructivism.

particular standpoint, and so on. Perhaps, at some point, these further beliefs will end up referencing the one we started out with. But, in this case, the account is circular (and vacuous).⁸⁵ There may be more that can be said here on behalf of such a view. But the prospects look dim. To the extent that this kind of view offers any substantive account of the content of these standpoint beliefs it opens itself to the first kind of objection – viz., that such beliefs (covertly) make reference to some belief-independent ethical truth. So far, we have been supposing that the stances that make up an agent’s standpoint are something like beliefs. But, what if they are some kind of non-belief-like state? How should we understand the states of mind that these basic ethical stances embody?

Suppose now that standpoints are composed of stances that embody some non-belief-like state of mind. Considering some of the motivations that constructivists appeal to, one might assume that these basic stances embody a form of pro-attitude, like desires, with a different direction of fit – one according to which we aim to conform the world to the content of our attitudes. But if one is able to work out the constitutive relations amongst this class of non-belief-like pro-attitudes, one will have arguably succeeded in one of the projects that expressivists are engaged in.

One of the big challenges for expressivists is to explain how ethical discourse exhibits standard logical inference patterns when, on their view, ethical language is not fact-stating. Standard logical inference is truth preserving. But expressivists don’t think that ethical language is used to express truth-evaluable content. So, expressivists must come up with an alternative “practical logic” which accounts for why it is

⁸⁵ See Ridge (forthcoming) for a similar objection.

nonetheless legitimate to use ethical language in these ways. Although different proposals have been made and developed, they remain extremely controversial.⁸⁶

Constructivism is usually understood as a version of cognitivism. This is because it makes ethical judgment, at some level, a species of belief. As I have presented it thus far, the standpoint constructivist takes it that sincerely uttered ethical statements generally express beliefs – viz., ones about what follows from within a particular agent’s ethical standpoint. They have truth-evaluable content and, as such, should figure unproblematically in contexts which require such content (belief reports, truth-preserving inferences, etc.). This might create the impression that constructivism will be immune to the kinds of objections that expressivists face. But if the stances that constitute these standpoints are themselves non-belief-like, it would appear as if the constructivist is saddled with the same project that the expressivist is, at least at the level of the lower-tiered ethical stances.

Constructivists must explain what they mean when they say that certain stances *must follow* from others in a way that does not rely on the notion of truth-preservation. We might say that the project of explaining this involves devising a kind of *practical logic* – in contrast with standard logic, which explains inference on the basis of truth-preservation.⁸⁷ If certain non-belief-like stances are supposed to follow from within a particular standpoint constituted by other non-belief-like stances, we must suppose that the relations amongst these stances constitute a structure with its own non-truth-preserving practical logic. For example, one might characterize this structure in terms

⁸⁶ For an extended critique of expressivist attempts to meet this challenge, see Schroeder (2008).

⁸⁷ It is perhaps for this reason that Gibbard (1999) reads Korsgaard’s constructivism as a species of expressivism. He takes her to be engaged in a project of working out a “practical logic”. I take this expression from him.

of consistency, or coherence, in attitude. But such a project will take considerable work if it is to succeed in accounting for all of the standard patterns of inference exhibited in everyday ethical discourse. What's more, whether it succeeds or not will likely be extremely controversial. However, it is only once this logic is worked out that a constructivist will be able to explain why a particular ethical judgment is true or not. If this is indeed the way to understand the standpoint constructivist's apparatus, one might object that the view does no better, or even perhaps worse than expressivism.⁸⁸ In fact, an expressivist might argue that the constructivist ought to avoid the complication that comes with taking ethical language to express beliefs and simply abandon cognitivism altogether; instead, she should take ethical statements to directly express the non-belief-like states of mind that constitute the various ethical standpoints. This would in effect make the view a species of expressivism.⁸⁹

This is a line that has become increasingly popular recently with several defenders of expressivism arguing, in various ways, that constructivism is best understood as a species of, or supplement to, their view – see, e.g., Gibbard (1999), Lenman (2010), Ridge (forthcoming). It has even been encouraged by some constructivists – e.g., Korsgaard (2003) and, to a lesser extent, Street (2008), (2010).⁹⁰

⁸⁸ For example, it may be objected that it does worse at accommodating the putative necessary connection between ethical judgments and motivation. If the constructivist maintains that ethical judgments are beliefs what follows from within a class of non-belief-like stances, she will still have a problem explaining the allegedly *necessary* connection between these “distinct existences”. Expressivists claim to avoid this objection by doing away with the idea that ethical judgments express beliefs; rather, they claim that an ethical judgment directly expresses a state-of-mind that is, perhaps amongst other things, intrinsically motivating.

⁸⁹ Street appears to flirt with this idea. See, e.g., Street (2008), 230; (2010), 375ff.

⁹⁰ Korsgaard (2003): “Expressivism, I believe, is like realism also true after all, and also in a way that makes it boring” (p. 122n49).

But if constructivism is indeed best understood in this way it arguably loses in importance as a challenger to realism. The expressivist challenge to realism is both familiar and fairly well understood. Moreover, a constructivist-expressivism would represent an even less palatable version of this challenge than the existing “quasi-realist” versions advanced by expressivists like Blackburn and Gibbard.⁹¹ Not only would a defender of constructivist-expressivism be giving up on cognitivism, she would also be giving up on even the appearance that ethical discourse is stance-independent. Quasi-realist expressivists are at least concerned to accommodate realist ways of talking and thinking about objectivity in ethics. Someone who defends this kind of constructivism could not claim to get you everything the realist wants without the problems. Such a view arguably fails to get you *anything* the realist wants. Although such a view might represent an interesting internal challenge for expressivists, it is not the kind of challenger that I am setting out to evaluate in this work. But the only apparent way around this objection risks making constructivism into a version of another well-known metaethical position.

Any version of constructivism that characterizes ethical standpoints in terms of non-belief-like stances will have a problem distinguishing itself as a free-standing view. Suppose that the constructivist denies that an ethical standpoint involves a level of structure that would require a practical logic. For example, a constructivist might deny that the truth of one’s ethical judgments follows in any interesting way from the relations which an agent’s ethical stances bear to one another – though the truth of such judgments would still follow from facts about these stances taken together, along

⁹¹ See, e.g., Blackburn (1984), 197; Gibbard (2003), 18ff.

with facts about the world. This would position the view closer to a form of simple *subjectivism* – e.g., a view that takes moral judgments to express beliefs about which actions would maximally satisfy an agent’s actual desires.

One might also describe this desire-satisfaction view as providing a similarly ordered account of ethical judgments. At the level of the agent’s ethical standpoint there are desires, a kind of non-belief-like stance; at the level of ethical judgment there are beliefs about these desires. Furthermore, one might argue that such a view, like a standpoint constructivism, is distinct from the family of response-dependence theories. Ethical judgment on the desire-satisfaction view is not about how certain subjects would respond; rather, it is about whether an action satisfies one’s actual desires, and how many.

Importantly, however, the desire-satisfaction theorist does not take ethical judgments to be about what *follows* amongst an agent’s desires; our ethical judgments, on this view, are not beliefs about the relations that an agent’s desires bear to one another. Rather, they express beliefs about how many desires on balance would be satisfied, or frustrated, by a particular course of action. Such a view arguably does not require any logic. Although an agent’s set of desires may exhibit some structure – e.g., with some desires taking other desires as their objects, or some desires being more general and others specifications of them – it does not involve a level of sophistication that would support “entailment” claims. Consequently, the challenges associated with expressivism do not arise.

The desire-satisfaction theory provides a model for how a constructivist might avoid the kind of difficulties associated with the expressivist’s project. Of course, the

problem with this approach is that it requires that the constructivist explain how her view represents a novel and interesting advance on the desire-satisfaction theory, or other common versions of simple subjectivism. One might have thought that the extent to which constructivism represents an improvement on such theories is the extent to which the view incorporates structure at the level of the agent's ethical standpoint.⁹²

There would appear to be a dilemma for a constructivist who insists that ethical standpoints are composed of non-belief-like stances. Either these stances combine to create a structure within which some stances may be said to follow from others, in which case the view involves the same difficulties that expressivism does. Or a standpoint is to be understood as a mere collection of stances without any sophisticated structure, in which case the view starts to look like a version of simple subjectivism, with all the virtues and vices that such views carry.

2.6 Reframing the Challenge

The prospect of a freestanding metaethical constructivism is looking dim. The standard proceduralist characterizations give the appearance that constructivism is best understood as a version of a response-dependence theory. This might make the

⁹² For example, Street (2008) describes the difference in this way:

Yet normative judgments are different from *desires* in virtue of the kinds of constitutive involvements I've been sketching. For example, whereas *taking oneself to have a reason* to live constitutively involves *taking oneself to have reason* to undergo the leg amputation that one knows is necessary, the attitude of desire is characterized by no analogous constitutive involvement: one can *desire* to live while having no *desire* whatsoever to under the leg amputation. (pp. 230-1).

standpoint characterization appear more promising. But if an ethical standpoint is constituted by beliefs, constructivism either folds into realism or turns out to be vacuous. If it is constituted by non-belief-like stances, it is best interpreted as a species of, or supplement to, expressivism or a version of a simple subjectivism. Unless constructivism can be shown to represent an advance on one of these alternatives, the view would appear to lack in motivation. Nothing I've said thus far rules out this possibility. But even if constructivism represented an improvement on response-dependence theories or expressivism or simple subjectivism, it may still fail to represent any new or interesting challenge to realism. One would have to show that constructivism improves on these other views in ways that also makes for a more formidable challenge to realism.⁹³ Perhaps the more promising option would be to give up on the claim that constructivism represents a free-standing metaethical view and argue that the challenge for realism is of a different sort.

If the arguments in this chapter are sound, it would not appear as if any version of constructivism – formal or substantive – counts as a free-standing metaethical alternative to already familiar views. Nevertheless, I think constructivist arguments still pose an interesting, though somewhat more *restricted*, challenge to realist commitments.

Specifically, constructivists may be seen as targeting realist views of *ethical objectivity* – a mere part of the realist's total metaethical package but, nevertheless, an essential one. Realists might not be alone in accepting conditions (3) and (4). For example, representative quietists, error-theorists, and quasi-realist expressivists have

⁹³ This is because each of these views has their own motivations and challenges. How a view stacks up to realism is only one metric of evaluation.

each laid claim to the view of ethical objectivity expressed by these two conditions. But, as I have argued, these conditions also represent constitutive commitments of any (robustly) realist view in ethics.

There are two ways in which constructivists might challenge the realist's view of ethical objectivity, one corresponding to each of the two conditions. Formal constructivists take aim at condition (3). They reject the realist's view of the *nature* of ethical objectivity. If they are correct, all ethical truth is stance-dependent. One of the questions that we must ask then is whether we lose anything important if they are right. I address this question in Chapter 3. Substantive constructivists take aim at condition (4). Although they accept that some part of the ethical enjoys the kind of objectivity expressed in (3), they reject the realist's view on the *extent* of this kind of objectivity. If they are correct, the truth of at least some of our moral principles is stance-dependent. Here one of the questions that we might ask is why we should think that the realist's notion of objectivity should extend to as much of the ethical domain as (4) would imply. Mightn't there be other ethical considerations that undermine (4)? I address this question in Chapters 4 and 5.

CHAPTER 3

FORMAL CONSTRUCTIVISM AND CONSIDERED MORAL JUDGMENTS

3.1 Introduction

Although current consensus has it that formal views offer the only coherent and defensible constructivist challenge to a stance-independent realism, I argue in this chapter that these views are also the most vulnerable to challenge from within ethical theory. This is because, as *formal* views, they lack the resources to account for certain widely and deeply held substantive moral convictions – ones, it turns out, that formal constructivists themselves acknowledge and accept. What's more, these constructivists accept a method of theory-building that commits them to providing some account of these judgments. They must show that a formal constructivism either supports these judgments or provides some way of revising or explaining them away.

Every defender of constructivism, both formal and substantive, assigns *reflective equilibrium* some role in their theory building.⁹⁴ Very roughly, this method involves comparing, revising, and pruning our ethical judgments with the aim of reaching a coherence, or equilibrium, amongst them. This process of reflection starts with our *considered judgments*. These are deeply held ethical judgments at all levels of generality, both judgments about particular cases and theoretical considerations that might bear on them. From these starting points, we then proceed to work back and

⁹⁴ The role may vary. Some assign reflective equilibrium a constitutive role. On this view, ethical truth just is that a judgment stands in reflective equilibrium with an agent's other judgments. Others assign it a heuristic role. On this kind of view, that a judgment stands in reflective equilibrium provides good reason to believe that it is true. See Street (2008) for discussion of these differences.

forth between general principles and particular judgments, modifying them according to norms of consistency and coherence, until we reach a state of equilibrium in which considered judgments at all levels of generality exhibit an acceptable degree of support for one another.

Since constructivists all accept this method, it is fair to ask whether the specific versions of constructivism they defend can offer support to our considered judgments in reflective equilibrium. I argue, however, that despite defenders' best efforts formal versions of constructivism cannot. This deficiency, together with the importance that constructivists themselves assign these judgments, makes for a serious objection to formal views. What's more, since realism provides a way of accounting for them, one may speak of a comparative advantage for realists.

In section 2, I explain in more detail what makes a version of constructivism formal and what considerations are taken to support this kind of view. In section 3, I present a challenge to formal constructivism from within ethical theory. In section 4, I draw a further distinction between two types of formal views. Although both types turn out to be compatible (in some sense) with the considered judgment at the heart of my challenge, only one bears any chance of being able to justify it. In section 5, I argue however that even this more promising type of formal constructivism fails to account for the judgment at issue. In sections 6 and 7, I entertain rejoinders. In section 8, I conclude with a few remarks on how the results of this chapter motivate the remaining two.

3.2 What is Formal Constructivism?

The distinction between formal and substantive versions of constructivism is one about how we are to characterize an ethical standpoint. Specifically, it concerns what kind of materials we may appeal to in these characterizations. Again, a version of constructivism is *formal* if it characterizes an ethical standpoint by appeal to strictly formal norms (e.g., consistency, coherence, instrumentality). Formal norms apply to the structure of an agent's ethical judgments taken together. They specify the way in which judgments, irrespective of their particular contents, must hang together in order to make a standpoint what it is. In other words, these norms do not presuppose the truth of any ethical judgment – general or particular. Rather, according to such views, an ethical judgment is true just in case (and because) it follows from within an ethical standpoint, where this standpoint is characterized formally. On a formal view, there are no ethical judgments that are excluded from this requirement.

Formal versions of constructivism are a recent development. They have been proposed in response to objections against standard substantive presentations of the view (e.g., those one finds in the work of John Rawls, T.M. Scanlon, and Christine Korsgaard). Since substantive views do nothing to justify the class of basic, status-conferring ethical judgments (i.e., they do nothing to explain why this class of judgments should play this kind of foundational role), they are compatible with a stance-independent account of the truth of these judgments. If this is correct, it would appear as if these constructivists accept (3) and with it at least part of the realist's account of the nature of ethical objectivity.

Some constructivists, like Sharon Street, appear to think that only a *thoroughgoing* constructivism could challenge realism – that is, one according to which *all* ethical truth is entailed from within the ethical standpoint. Formal versions of constructivism are thoroughgoing and, hence, supposed to avoid this objection. On these accounts, there is no privileged class of ethical judgments; rather, every judgment is subject to scrutiny from within the standpoint. In slogan form: ethical truth is constructed “all the way down”. Hence, these constructivists reject the realist’s condition (3).

Formal constructivists claim that someone who accepts their view does better at accommodating other deeply held first-order ethical judgments. The defense is holistic, meaning that formal constructivism must be evaluated against a host of considerations. The claim is not that the view succeeds in accommodating every consideration; rather it is that the view does the best job on balance of any relevant alternative. In other words, the claim is that formal constructivism, not a stance-independent realism, stands in reflective equilibrium with our considered judgments.

3.3 An Internal Challenge for Formal Constructivism

Are there particular ethical claims that formal constructivism fails to accommodate? And, if so, does this failure constitute a significant cost for the view on the whole? I argue *yes* and *yes*.

Consider the moral claim that *torturing children for fun would be wrong even if no one ever took it to be such*. This claim arguably expresses a considered moral

judgment, not only one that many would accept but also one that any good moral theory should be in a position to *justify* to some extent. That is, the success of a moral theory will depend in part on whether it is able to provide some good reason for thinking that this claim is true.

At the very least, an ethical view must be compatible with this claim (henceforth, “the torture claim”) if it is to be plausible. That is, it should provide no good reason for thinking that this claim is false. This point is not trivial. Some views *are* incompatible with it. Consider, for example, the view that an action is wrong just in case we all decide that it is. Let us call this view *morality by consensus*. It can allow that torturing children for fun is wrong and might even provide resources for justifying this claim. Perhaps under normal circumstances we would all decide that this kind of action is wrong. But it is incompatible with the last clause – *even if no one ever took it to be such*. According to morality by consensus, torturing children for fun is wrong *if and only if* we decide that it is. This means that it would not be wrong if we all decided that this kind of torture was permissible. Results like this lead many to reject this kind of view as objectionably relativistic and deeply implausible. Other considerations might count in favor of such a view, but it is hard to see these competing considerations ever outweighing the costs. Again, if formal constructivism is to stand a chance it must at least be compatible with the torture claim. But even if it is compatible, one may still ask how well it does at justifying this claim.

Although formal constructivism *is* compatible with the torture claim, I argue that it lacks the resources to justify it. This result represents a serious defect of the view – one that undermines the claim that constructivism, on balance, is a superior

alternative to realism, or other views that are in part constituted by the claim that ethical truth is stance-independent.⁹⁵ In fact, realists would appear to enjoy a comparative advantage. According to realism, whether or not torturing children for fun is wrong will depend on features of the world; it does not depend on our taking such actions to be wrong.

Is this a fair challenge for the constructivist? One might worry that my challenge begs the question. If the constructivist thesis entails that the torture claim is false, my challenge would merely assert what constructivists deny. This would not represent a fair or productive way of evaluating the view. But as it turns out formal constructivism does not entail that the torture claim is false. In fact, defenders of formal constructivism recognize this kind of objection and appear concerned to address it.

According to formal constructivism, our taking something to be wrong is neither necessary nor sufficient for its being wrong; rather, in order to be wrong, the judgment that it is so must follow from within an agent's *moral standpoint*.⁹⁶ This standpoint is constituted by certain (formal) norms together with the other moral judgments that an agent holds. Arguably this means that the constructivist may maintain that a particular judgment follows from within an ethical standpoint without someone's *actually* making this judgment.⁹⁷ The judgments entailed from within this

⁹⁵ Here I have in mind a quasi-realist expressivism like the one defended by Simon Blackburn (1984), (1993) or Allan Gibbard (2003), as well as the kind of "realist" quietism one finds presented in the work of Ronald Dworkin (1996) and Thomas Nagel (1989).

⁹⁶ Street (2010), 367n. 16.

⁹⁷ I specify *actually* here because it turns out that hypothetical judgment is trickier. Could the judgment count as true for a formal constructivist even if no one would make this judgment under idealized, hypothetical circumstances? Here it really depends on the idealization one employs. If one specifies the idealization so that one has "whatever it takes" for one's judgment to line up with what is entailed from

ethical standpoint are the ones we *ought to* maintain. But we can be mistaken. Perhaps on this view it is even possible that everyone is always mistaken about certain judgments.⁹⁸ All the constructivist is committed to saying is that the correctness of this judgment depends on a person's taking something or other to be wrong or right or permissible; not that this person takes the very action in question to be wrong. So it looks as if constructivism is at least compatible with the torture claim.

The next question is whether a formal constructivism has the resources to justify this claim. In what follows, I will argue that it cannot. My evaluation will involve looking at different kinds of formal constructivism and the various responses that their defenders have proposed to the kind of objection I am presenting.

3.4 Two Kinds of Formal Constructivism

Formal constructivism comes in two flavors. Some defenders of formal constructivism claim that no matter what judgments one starts out with there will necessarily be a convergence on a single class of correct ethical judgments.⁹⁹ Others deny this.¹⁰⁰

within the standpoint, it turns out that such a judgment would not count as true for the constructivist. Note, however, that idealized judgment does not constitute truth for the constructivist. Formal constructivism is formulated in terms of standpoints in order to avoid the problems with a hypothetical proceduralism. Although a defender of the standpoint characterization claims that ethical truth is existentially dependent on an agent's making some ethical judgments or other, she denies that it is constituted by the judgments of a suitable described agent under special conditions. This means that the truth of particular ethical judgment may outstrip an agent's ever making the judgment, actually or hypothetically.

⁹⁸ There must be a limit on this, however, since constructivists level skeptical arguments against realists. For example Street (2006), (2008), (2009a) argues against realism on the grounds that it fails to provide any good explanation of why our actual ethical judgments turned out to be the true ones. It does not appear as if the constructivist would want to admit that her own view was compatible with everyone's being mistaken about all of our ethical judgments.

⁹⁹ See James (2007). Alternatively, we might say that there will be a convergence on a core of ethical truths, even if there are differences at the periphery.

Although they allow that there may be convergence, they deny that it is necessary. On their view, correct ethical judgment is always a function of the judgments one starts out with. Let us follow recent convention and call this second kind of view a formal *Humean* constructivism and the first kind a formal *Kantian* constructivism.¹⁰¹ Of the two types, the Kantian view appears to present the better chance of justifying the torture claim and, consequently, the stronger challenge to realism.

A formal Humean constructivism is compatible with the torture claim. But it does not provide any support for it. That is, it provides no reason to think that the torture claim is true. Again, according to this view, what is entailed by an ethical standpoint depends on what judgments one starts out with. Although the formal, constitutive commitments of an ethical standpoint might constrain outcomes to some extent, as *formal* commitments they will only do so much. They do not contribute anything of content to the outcome. Presumably this means that two agents who start out with radically divergent judgments might end up with some divergence in the class of output judgments.

Consider the case of what Street (2009) calls an *ideally-coherent eccentric* [ICE]. ICEs are hypothetical characters who possess two main features:

¹⁰⁰ See Street (2008), (2010).

¹⁰¹ NB I am following Street's terminology here. However, it is important to note that on this usage the class of formal Kantian views is quite narrow. For example, it does not cover the kind of *Kantian constructivism* that Rawls (1980) discusses. This view, as I have noted earlier, would count as a substantive, not formal, constructivism. There is some interpretive dispute as to whether the view in Korsgaard (1996), (2003) ought to count as a formal Kantian view. Street (2008), (2010) interprets Korsgaard in this way. But I think Korsgaard's view is better seen as a substantive version of constructivism. For discussion, see Chapter 1, sections 5 and 6.

First, they accept some value that is utterly unheard of, morally repugnant, or both. Second, their acceptance of this value coheres perfectly, as a logical and instrumental matter, with all of their other values in combination with the non-normative facts.¹⁰²

These characters are marched out as counterexamples to the claim that merely formal criteria, like those described in the second feature, are sufficient to generate an account of reasons, values, or morality that coheres with our considered judgments.¹⁰³

A Humean might claim that in a society like ours, one with the typical beliefs and attitudes that people have, it is likely that an agent's moral standpoint will entail the torture claim. But in the case of ICEs, it is at least conceivable that it will not. The result, according to the Humean view, would be that the torture claim was not true *for them*. But importantly the torture claim is not supposed to be true for us, or people like us. It is supposed to be true for everyone, regardless of the judgments they actually hold. We do not simply withdraw our claim that such people morally ought not to torture if we find out that such actions are compatible with the judgments entailed from within their own respective ethical standpoints. Let us try to identify exactly what claim these counterexamples are meant to target.

¹⁰² Street (2009), 273.

¹⁰³ For example, Gilbert Harman's (1975) Hitler, whose moral repugnance is infamous, and Allan Gibbard's (1999) Caligula, who "aims solely to maximize the suffering of others", are stipulated to possess both of these features. Here it is more appropriate to speak of an *idealized* Hitler. For, as Harman himself notes, the actual Hitler in all likelihood did not satisfy the second feature that Street describes. The same should be said of the Caligula case.

Although a formal Humean view is compatible with the torture claim, it is incompatible with a closely related claim that many would also accept. Consider

Torture Plus:

It's not just that torturing children for fun would be wrong even if no one took it be such; rather that it would be wrong regardless of whatever actual judgments one starts out with.

A formal Kantian constructivist can accept the second part of this claim, but a Humean cannot. On the Humean view, torture's wrongness is independent of what we actually think of it but it is *not* independent of the set of our actual moral judgments taken together. And if these other judgments are deviant enough, as in the case of ICEs, it is possible that torturing children fun turns out to be permissible. That is, even if the view does not make (in this case) *moral* truth dependent on our, or anyone's, actually *taking the particular action to be wrong*, it makes it dependent on one's other ethical judgments and allows that they may wield a considerable influence over whether it turns out that torturing children for fun is wrong. The result is that the Humean view is compatible with an objectionable relativism. This makes the view deeply implausible.

But is it fair to evaluate formal Humean constructivism against Torture Plus? If Torture Plus is true, Humean constructivism is false. If we merely stipulated Torture Plus, this would be unfair. But there is a quick argument from the torture claim to Torture Plus. If it is sound, it will provide independent reason to accept Torture Plus

and reject the Humean view. That is, Torture Plus is controversial but it can be shown to follow from the torture claim itself, a claim that all parties can agree on.

Torture Plus can be shown to follow from the simple torture claim, a considered judgment that we have independent reason to accept. Consider two counterfactuals.

C1: If I took torturing children for fun to be permissible, it would be.

C2: There is a set of ethical judgments very different from the ones I actually hold, such that if I started out with this set, torturing children for fun would turn out to be permissible.

I will argue that we should reject both counterfactuals. This is more controversial in the case of C2 than C1. For example, a defender of the Humean view can reject C1 but she cannot reject C2. Why is this? The Humean allows that the formal norms constituting an ethical standpoint may constrain what judgments ultimately count as true. So, just holding a particular judgment doesn't suffice for its truth. But the Humean *does* make ethical truth contingent on one's actual judgments taken together. This commits the Humean to C2. But if one rejects C1 there is also good reason to reject C2. Consider this little argument.

Suppose we reject C1. So there is at least one case where one's actually taking something to be permissible is not sufficient for its being so. But this case is not unique. According to formal constructivism, a judgment is never true merely in virtue

of the fact that someone actually makes it. Rather, it is true if and only if it follows from within the agent's ethical standpoint as such, where this standpoint is constituted by the agent's actual judgments taken together in accordance with certain formal norms (e.g., consistency, coherence, instrumentality). But if any particular case of taking something to be wrong does not suffice for its being wrong, why should we think that our actual judgments taken together suffice to make it the case that any particular action is wrong? If my actually taking torture to be permissible does not suffice to make it permissible, and my taking any other particular action to be permissible does not suffice to make it permissible, then it is hard to see how these judgments taken together should suffice. Even if all of my particular judgments were very different, together they still would not make it the case that torturing children for fun was permissible. So, we should reject C2 as well as C1. Since a Humean view commits one to C2, such a view would be mistaken. In this respect, versions of formal constructivism that include a claim about necessary convergence appear to have the upper hand.

A defender of a Kantian view might claim that the torture claim *is* supported by a judgment that falls into the convergence class. That is, no matter what judgments one starts out with, an ethical standpoint is guaranteed to entail certain outcomes as correct. Furthermore one might claim that, on this view, it turns out that the torture claim is amongst these (alternatively, there is some ethical judgment in this class that supports the torture claim). If all this were true, a defender of a formal constructivism could claim to neutralize the challenge I have presented. But, as I will soon argue, it is not true; or rather there is no good reason to think that it is true. Insofar as a

constructivism remains formal, it lacks the resources to fend off the challenge I have presented. It will not be able to justify the torture claim, or many other claims that match our considered judgments.

3.5 The Failure of Formal Kantian Views

The defender of formal Kantian constructivism makes a couple of different claims. Hence, there are different ways in which her view might fail. Let us distinguish two claims:

- (i) Formal constraints guarantee convergence on some class of correct ethical judgments.
- (ii) Formal constraints guarantee convergence on some class of correct ethical judgments that includes the torture claim itself (or at least others that support its truth).

One way in which the Kantian could fail would be if she could not justify the truth of (i). This claim defines what it is to be a formal Kantian constructivism. If this claim were false, there would be no alternative to Humean views to speak of. What's more, claim (ii) entails (i). So, if (i) were false, (ii) would also be false. If formal Kantian constructivism fails with respect to (i), the alternative is for the constructivist to adopt a formal Humean view. And, as I have just argued, such views are especially vulnerable to the challenge I am presenting.

Another way in which the Kantian could fail would be if she could justify the truth of (i) but not (ii). This result might mark an improvement over Humean views to some extent. At the very least, it would show that there is a formal constructivist alternative to Humean views. But failure to justify the torture claim would still count as a serious defect of the view. The challenge for the Kantian is to provide some good reason for thinking that a purely formal characterization of an ethical standpoint provides any reason for maintaining both (i) and (ii).

As I will soon argue, a formal constructivist may succeed in defending (i), but she lacks the resources to defend (ii). This result generalizes. To the extent that the view fails to accommodate (ii) it will also fail to accommodate many of our particular moral judgments.¹⁰⁴ Hence, the Kantian may present a genuine alternative to formal Humean views. But she fails to accommodate our considered moral judgments. Let us first consider the best formal constructivist defense of (i) and then ask whether it might be extended to support (ii).

A formal constructivist might try to defend (i) by appealing to a *norm of universality*. This kind of norm has appeared plausible to many, especially those working in the Kantian tradition. But we can take universality in different ways. On the one hand, we might take it to affect the scope of a moral judgment. For example, a norm of universality might require that any correct moral judgment be derivable from principles that one can take to apply to everyone in the same way. On the other hand, we might take it to require that a judgment be derivable from principles that would be

¹⁰⁴ *Many* but not all, since it might succeed in securing the truth of very general judgments.

accepted by everyone.¹⁰⁵ But it is not clear that either norm would help the formal convergence constructivist defend (i) and (ii). The first does not secure convergence across agents. The latter does, but it does not appear to be formal.

The first candidate constrains the scope of moral judgments. In order to count as genuinely engaging the moral standpoint one's judgments must apply universally. According to such a norm, if one takes some action to be wrong (for example) one commits oneself to applying the same standard to all particular judgments involving similar circumstances, regardless of whom the judgment concerns. Failure to commit oneself in this way would mean that one would not genuinely count as *taking an action to be wrong*. Such a norm would count (along with instrumentality, coherence, consistency, etc.) as a constitutive commitment of the moral standpoint as such.

But such a norm does not guarantee convergence across agents. That is, it does not guarantee that different agents will be committed to the same class of judgments; rather, it only commits a particular agent to applying her own judgments in a particular way. It will constrain outcomes to some extent. For example, if I judge that what Tammo does in circumstances *C* is wrong, I am committed to judging that Timmy is wrong when he performs the same action type in circumstances qualitatively indistinguishable from *C*.¹⁰⁶ But this kind of universality does not guarantee the relevant sense of convergence across agents. For example, if Gonzo judges that what Tammo does in circumstances *C* is permissible, he is committed to judging that it is also permissible for Timmy to act this way in qualitatively indistinguishable

¹⁰⁵ Thanks to Nicholas Sturgeon for helping me see this distinction.

¹⁰⁶ I am assuming that we might include a description of the agent's intentions and other relevant features of an agent in *C*.

circumstances. In short: this norm of universality will yield different sets of outcomes depending on the judgments one starts out with. It does not generate a single class of correct ethical judgments regardless of an agent's starting points. If a formal constructivism is to guarantee convergence in outcome (i.e., in the ethical judgments that every particular agent is committed to), it must include a stronger norm.

The second candidate norm involves an agreement constraint. It requires that the judgments one make conform to principles that would be accepted by all agents to the extent that they meet certain other conditions.¹⁰⁷ Alternatively, we might understand this requirement in terms of universal commitment across agents. That *all* agents would accept (or be committed to) principles grounding a judgment is to say that there would be convergence on the truth of this judgment across agents. In other words, a condition for my judgment's counting as correct would be that everyone else would make (or be committed to) the same judgment. The problem with this kind of norm is that it is not formal. This is because the norm requires that a relation obtain between an agent's stances and those held by others; it is external to an individual standpoint.

A formal norm specifies the relations that must obtain amongst a set of stances so that they hang together in the right way to constitute a standpoint. Arguably these relations are all *internal* to the set of judgments that an agent makes. That is, they only involve relations that obtain amongst the agent's own judgments. But this second formulation of universality relates one agent's set of judgments to the judgments that

¹⁰⁷ There is also the related problem of cashing out "certain other conditions" in a way that does not rely on stance-independent notion of ethical truth. I set this worry aside for now and focus on what is arguably a deeper objection to a *formal* norm of universality. But I will revisit it later in the discussion.

other agents make (or those that they are committed to by their own standpoints). Such a norm involves *external* relations.

It sounds odd to claim that a structural requirement for *my* set of judgments' constituting a standpoint is that they bear certain relations to the judgments that other agents make. Compare the case of belief. It would be odd to claim that one of the constitutive commitments of *belief* is that other people share this attitude with me. Although we may speak of the truth or, more controversially, the content of one's belief being determined by external features of the world (including, perhaps, what other people think), it is odd to think that my belief's genuinely counting as the type of attitude it is depends on such external features.¹⁰⁸ Similarly, it is odd to think that my moral judgment's genuinely counting as the kind of attitude it is depends on such external features, even if the content or truth of this judgment depended on them in some way.¹⁰⁹ Rather, it would appear as if a formal norm must only involve internal relations. In this case, universality still doesn't get the formal constructivist (i) because the norm involves external relations and a norm that involves external relations is not *formal*. But even if this is wrong (i.e., it turns out that a formal norm may involve external relations), this particular norm will not do the job. The second kind of universality norm may secure (i). But it still won't be enough to justify (ii).

¹⁰⁸ I do not wish to take a stand on these controversial topics. My point is simply that these kinds of external views (i.e., ones about content and truth-conditions) are open to debate. An external view of the formal constitutive requirements for moral judgment, by contrast, is not.

¹⁰⁹ What's more, if constitutive norms included such external features, a formal constructivist would undermine some of the claims she takes to motivate her view. Specifically, it would undermine the constructivist's claim that her view enjoys an epistemological advantage over realism. For example, what reason do we have to believe that our moral judgments in fact meet this requirement? How could we ever know this?

Suppose that it is legitimate to speak of formal norms that include external relations. In this case, formal constructivists might characterize a moral standpoint in part by appeal to a universal norm of agreement or commitment amongst others. Is there now some reason to think that the torture claim will be a member of the convergence class entailed by such a standpoint, or supported by other members that are? The answer would appear to be *no*.

Let us assume that such a norm yields necessary convergence on some class of judgments. That is, no matter what judgments you start off with, the norm of universality leads to convergence on some non-empty class of correct judgments.¹¹⁰ Will this class include or support the torture claim? There is no good reason to think it will.

The second version of the universality norm does not support the torture claim. The problem is that this norm appears consistent with the contradictory of the torture claim: viz. torturing children for fun is permissible even if no one takes it to be such.¹¹¹ All that the norm requires is that a judgment conform to principles that would be accepted by all agents who meet certain other conditions (or principles that all such agents would be committed to). Until we specify these conditions, there doesn't seem to be any reason to think that all other agents wouldn't agree to this contradictory principle. Surely they *could* all agree to it (or be committed to it). If they did, this would be a case where the ethical standpoint entails that the torture claim is false. This would, of course, be a bad result for the defender of a formal Kantian constructivism.

¹¹⁰ Presumably there are at least some moral judgments that satisfy this norm. That is, the standpoint in part constituted by universality is not *empty*; it entails that some judgments are correct.

¹¹¹ Assumption: not (torture claim) = torturing children for fun is permissible.

There are ways of supplementing the norm so that it rules out such a result. But the obvious candidates involve appeal to substantive, ethical considerations. Consider, for example, the norm that a judgment must conform to a principle that would be agreed upon by all agents who are concerned to respect humanity in their own person and all others.¹¹² Torturing children for fun (I think it is safe to assume) does not respect their humanity. Hence, there is good reason to believe that any set of judgments that conform to this supplemented norm of universality will rule out torturing children for fun. But the supplemented norm involves appeal to a substantive ethical consideration – viz., that humanity is valuable and worthy of respect. This means that any norm that rests on such a consideration will itself be substantive – *not* formal.

Hence, even if we allow that formal views may appeal to universality to guarantee some convergence, appeal to this norm is not sufficient to guarantee the truth of the torture claim. Such a norm – though compatible with this claim – is equally compatible with its contradictory. This is a troubling result for the formal constructivist.

¹¹² Compare the way that Kantians supplement the Formula of Universal Law (FUL) with the Formula of Humanity (FH). Many worry that FUL – like the formal universality norm I’ve described – is insufficient to get us *moral* outcomes. FUL says “Act only according to that maxim through which you can at the same time will that it should become a universal law” (G 4:421). But, for example, it is not clear how this norm alone would rule out the contradictory of the torture claim. Surely it seems possible to will a universal law that would allow torturing children for fun. It is hard to see what kind of practical inconsistency would be involved in so willing. In response to this kind of worry, some defenders of Kant look to other formulations of the law of morality (i.e., *the categorical imperative*) that one finds in Kant’s writings. This includes FH, amongst others. FH says “So act that you use humanity, whether in your own person or that of another, always at the same time as an end, never merely as a means” (G 4:429). I cannot evaluate the success of this kind of move. But it does appear as if FH stands a better chance of securing the right extension. For example, torturing children for fun does not use their humanity as an end in itself. The problem with an appeal to FH, however, is that it involves appeal to a substantive ethical consideration – viz., that humanity is valuable. But a *formal* constructivism cannot appeal to such a substantive value without undermining itself.

If formal constructivism is to claim an advantage, on balance, over realism (or other stance-independence views), it must not only be compatible with the torture claim; it must also *justify* it to some extent. Of the two kinds of formal views, only those that guarantee convergence on a class of correct ethical judgments stand to have a chance. Humean views reject necessary convergence. According to these views, whether the torture claim turns out to be correct is contingent on the judgments that an agent starts out with. The result is that in the case of certain idealized eccentrics the torture claim does not turn out to be correct.¹¹³ Kantian views, by contrast, state that no matter what judgments one starts out with there will necessarily be a convergence on a single class of correct ethical judgments. If the torture claim, or some other claim that supported it, could be shown to belong to this convergence class, defenders of this kind of formal constructivism could claim to meet our challenge. But, as I have just argued, there does not appear to be a merely formal norm that provides reason to think that the torture claim is true. The best candidate norms – e.g., norms of universality – are equally compatible with this claim’s contradictory. Standard supplements to this kind of norm appeal to substantive ethical judgments and, as such, are not available to defenders of formal views. Hence, it would appear as if no version of formal constructivism can justify the torture claim. This marks a serious defect of the view. Hence, formal constructivism appears to do worse than realism in at least one

¹¹³ Even if we restrict ourselves to the actual world (one in which ICEs probably don’t exist) Humean views still have problems with the torture claim. Again, this claim states that torturing children for fun would be wrong *even if no one ever took it to be such*. As I argued in section III, someone who accepts this claim has good reason to reject C2. Importantly, C2 does not require postulating idealized eccentrics; rather, someone who accepts it thinks that the torture claim might have been false *for them*. Here I take it that the relevant possible worlds are still close ones. But the Humean is committed to C2. In fact, this counterfactual practically defines the view.

important respect. This undermines the constructivist's claim that her view is superior, on balance, to realism.

Defenders of formal constructivism acknowledge the kind of challenge I have been presenting. But some would disagree with the way that I have characterized their views. In fact, some would argue that formal constructivism, properly understood, avoids this objection. In the following sections, I entertain two possible responses and argue that they fail to neutralize the objection I have been presenting.

3.6 Is the Challenge Too Strong?

One might object that the challenge I am pressing does not just affect formal versions of constructivism; rather, some stance-independence views look to be equally deficient. If this is right, isn't the challenge too strong?

Formal constructivism isn't the only view that may have a problem justifying the torture claim. Some stance-independence views may have a problem, as well. Consider a simple *hedonistic act utilitarianism*.¹¹⁴ Very roughly, this view states that the rightness or wrongness of an action depends on the balance of pleasure over pain that it produces in the whole world relative to that of the available alternatives. The experience of pleasure and pain are not intentional stances that agents take. Rather, on this view, an action's wrongness will always be independent of an agent's taking it to

¹¹⁴ I take it that this view is compatible with a host of metaethical positions, including realism. However, as I will soon explain, it presents a stance-independent account of moral wrongness. Hence, it is incompatible with any formal constructivism.

be such.¹¹⁵ In the terms introduced in Chapter 1, the view provides a *weakly* – but not *strongly* – *mind dependent* account of moral truth. But if wrongness is contingent on facts about pleasure and pain, there is no guarantee that the torture claim will come out true. In fact, this is a standard objection against consequentialist views in general (of which hedonistic act utilitarianism is a species), i.e., that in principle they may yield results that diverge from our considered judgments about particular cases. If the challenge I am putting to the formal constructivist is merely a version of this objection, it would be unfair to appeal to it as a consideration that counts against it and in favor of stance-independent forms of realism, since not even all stance-independence views can justify it. Does my challenge apply equally to hedonistic utilitarianism? No.

The challenge I have presented for formal views is that they cannot justify the torture claim, even if they turn out to be compatible with its truth. Hedonistic act utilitarianism, by contrast, can provide some good reason for thinking that the torture claim is true, even if it cannot guarantee its truth. According to this view, the claim's truth depends on facts about pleasure and pain. It may turn out that once we investigate how such actions affect people we will see that they generally (if not always) produce a greater balance of pain over pleasure or less pleasure than the

¹¹⁵ Of course, what we take to be wrong may affect the balance of pleasure and pain. For example, if we take torturing children for fun to be wrong, we may be more pained when witnessing such acts than if we took no stance on the issue. This might then contribute to the overall balance of pain produced by the action. But even in this case facts about pleasure and pain would still be independent of the stances we hold. Our stances might bear a causal relation to what we experience as painful or pleasurable, but facts about pleasure and pain still do not reduce to facts about stances; pleasure and pain are (arguably) not stances.

alternative actions available to an agent.¹¹⁶ Indeed this seems likely to be the case since torture is painful! Whatever pleasure one may derive from subjecting others to this treatment is very likely to be outweighed by the pain it creates (at least in worlds that are anything like the actual world). So, not only does hedonistic act utilitarianism provide a clear approach to justifying the torture claim, it appears likely, in a world like ours, to justify it to some extent.

But can't even the Humean get you this much? Isn't it likely, in a world like ours, that agents start out with judgments that are likely to support the torture claim? It is true that the Humean may do as well as the utilitarian at justifying the claim that *torturing children for fun is wrong*. But this is not all that the torture claim involves. Recall, it says that this torture is wrong *even if no one ever took it to be such*. Neither the Humean nor the Kantian can justify this last bit.

Humeans are committed to the counterfactual claim that if one started out with very different ethical judgments, torturing children for fun might turn out to be permissible. As I argued earlier, someone who accepts the torture claim has reason to reject this counterfactual. That is, if you accept the torture claim you have good reason to accept Torture Plus and this second claim *is* incompatible with a formal Humean constructivism. Importantly, Torture Plus does not involve appeal to far away possible worlds. Someone who rejects it thinks that the torture claim might have been false *for them*. Here I take it that the relevant possible worlds are still close-by ones.

Kantians have a hard time establishing their view as a genuine alternative to Humean ones. But, even assuming that they succeed in this, they provide no reason to

¹¹⁶ The fact that this is generally but not always the case might support an indirect version over a simple act utilitarianism.

think that the torture claim is supported by the convergence class. In other words, on such a view, it may turn out that some actions are wrong even if no one ever took them to be such. But the formal Kantian constructivist cannot justify the claim that torturing children for fun is one of these.

So, the challenge I have presented is not too strong. It may pose a problem for formal versions of constructivism. But it poses no similar problem for stance-independence views, like hedonistic act utilitarianism. On such a view, there is no guarantee that the torture claim is true. But, given what we know about torture and its effects, hedonistic act utilitarianism provides good reason for thinking that the claim that torturing children for fun is wrong *even if no one took it to be such*. By contrast, neither version of constructivism can justify the entire claim. In worlds like ours, Humean views may justify the claim that such torture is wrong, but it cannot justify the claim that it would be wrong even if no one took it to be such. Formal Kantian views, by contrast, may justify the claim that *some* actions are wrong even if no one took them to be such, but they cannot justify the claim that torturing children for fun is one of these.

3.7 What if the Constructivist Rejects Moral Rationalism?

A formal constructivist might argue that all my challenge has shown is that constructivism has defects when construed as a view about morality. But ethical constructivism need not be comprehensive. For example, Rawls, Scanlon, and (arguably) Korsgaard each present non-comprehensive versions of constructivism. My

opponent might claim that the lesson to draw from this objection is that constructivists should restrict their views to practical reason. On this revised view, constructivists are engaged in a project of explaining reasons for action, not what it is for some action to be right or wrong, or some object or state-of-affairs to be good or bad. The torture claim, as formulated, is a moral claim. It is not a claim about an agent's reasons for action. This means that a failure to justify the torture claim should not be seen as a deficiency of constructivism per se.

Although this is a line that a formal constructivist could take (in fact, it has been proposed by Sharon Street),¹¹⁷ it significantly weakens the challenge that constructivism presents to the realist. First, it involves rejecting a popular doctrine that some constructivists have thought provided a comparative advantage to their view: viz., moral rationalism.¹¹⁸ Second, once moral rationalism has been rejected it turns out that there is little reason for realists to worry about constructivism.

According to the challenge I have presented, formal constructivism does not stand in reflective equilibrium with certain considered moral judgments. Specifically, it lacks the resources to justify the torture claim. In order to avoid this objection, it is not enough that a constructivist restrict her view to claims about reasons; she must also reject moral rationalism.

Moral rationalism, very roughly, is the view that moral requirements necessarily provide us with reasons for action. For example, according to this view, it follows that it is morally wrong for an agent to do something, only if that agent has

¹¹⁷ See Street (2009).

¹¹⁸ Substantive constructivists like Rawls and Korsgaard take the alleged truth of moral rationalism to count in favour of their views and against moral realism. I present and evaluate their arguments in Chapter 4.

reason not to do it. A formal constructivist might insist that her view is merely intended to provide an account of reasons for action, but unless she also rejects moral rationalism she will not succeed in avoiding the objection. According to a formal constructivist account of reasons, it might turn out that certain idealized eccentrics, e.g., an idealized Hitler or Caligula, have no reason not to torture children for fun. Together with moral rationalism, it would follow that it is *also* not wrong of them to torture. If it were, they'd have reason not to. So, it is not enough for the formal constructivist to restrict herself to reasons for action. If she is to avoid my challenge, she must also reject moral rationalism. But some (including prominent *substantive* constructivists) would take this rejection itself to represent a significant cost of the view.¹¹⁹

In response to this kind of worry, a formal constructivist might introduce a distinction between strong and weak versions of moral rationalism. Let us take *strong moral rationalism* to be the view that moral claims always provide us with all-things-considered reason for action. By contrast, we may understand *weak moral rationalism* as the view that moral claims always provide us with *some* reason for action. A formal constructivist might only reject the strong version, not the weaker one. In this way, she might claim to accommodate enough of what is appealing about rationalism without committing herself to the troubling conclusion that, e.g., it would not be wrong of Caligula to torture children for fun.

According to strong moral rationalism, if torturing children for fun is wrong, one has all-things-considered reason not to torture. But, per stipulation, we are

¹¹⁹ For the attractions of moral rationalism, see the discussion in Chapter 4, section 2.

supposing that an idealized Caligula does not have all-things-considered reason not to torture. Although the formal constructivist accepts this outcome about Caligula's reasons, she wants to avoid the conclusion that it is not wrong of Caligula to torture children for fun. So let us suppose that she rejects the strong thesis in favor of the weaker one. In this case, the torture claim may obtain without its also being the case that Caligula has an all-things-considered reason not to torture. All that the weak thesis requires in this case is that Caligula have *some* reason not to torture – and for all we have said it might turn out that he does.

The weak thesis allows that one's moral reasons come apart from one's all-things-considered reasons for acting. Of course, on this view, moral claims always provide one with *moral* reasons. But the weak thesis allows that moral reasons may fall short of what one has reason to do *full stop*. So, even if it is wrong of Caligula to torture children for fun, he may, nevertheless, still have all-things-considered reason to do this (perhaps because his moral reasons are silenced by other considerations, or outweighed on the whole).¹²⁰ This response works well enough for an opponent who grants that Caligula would still have moral reasons not to torture. But why must her opponent grant this?

If a formal constructivism is compatible with its being the case that Caligula has all-things-considered reason to torture, it would appear to be equally compatible with its being the case that Caligula has *no* reason not to torture. Why couldn't an opponent stipulate this, as well? On the Humean view, reasons of any kind are always

¹²⁰ Street (2009) has additional things to say to the person who would still take issue with this as result about Caligula's reasons for action, but I would like to set this worry aside. Again, my concern here is primarily the consequences that Street's formal Humean constructivism has as a view of the nature of morality.

a function of the attitudes and judgment an agent starts out with. Formal Kantian constructivists deny this. But, as I have argued, they lack the resources to back up this claim. So why not stipulate that the set of attitudes and judgments that Caligula starts out with, together with the non-normative facts, is such that his practical standpoint entails that he has *no* reason whatsoever not to torture? This would certainly make the case extreme, since it is hard to imagine there being no features of Caligula's ethical standpoint, together with the non-normative facts, that would entail his having at least *some* reason not to torture. But formal constructivists like Street are happy to admit far-flung thought-experiments.¹²¹ What's more, such a case would appear to be separated from the original presentation by degree, not kind. If this is correct, a formal constructivist must accept the challenge. But, in this case, we see that even weak moral rationalism is still too strong. On the weak view, if Caligula turns out to have no reason whatsoever not to torture, it still follows that it is not wrong of Caligula to do so. The initial challenge stands.

¹²¹ Street (2009):

I agree with those who think that the best overall theory of value must prove itself in its handling of hypothetical cases. Indeed, I would go farther than this: not only is it not silly to think about ideally coherent eccentrics, these characters haven't gotten anywhere near *enough* philosophical attention. In my view, ICEs are where the action is if we want to get clear on the relation between our attitudes, value, and the world... It is only in cases where ideal coherence is stipulated that we can be sure the two views will generate different consequences.... (p. 279)

There is a sense in which nothing hinges on what we say about ICEs, and there is a sense in which everything does. Nothing hinges on these cases in the sense that none of us will ever actually encounter an ICE. But everything hinges on these cases in the sense that our whole conception of ourselves and what we're doing when we engage in practical reasoning depends on the right thing to say about them. How should we understand the relationship between ourselves, and the world? When we face the deepest, toughest questions about how to conduct our lives, is the kind of fidelity we're after fidelity to something independent or is it fidelity to ourselves? ... we are forced by [the mere possibility of ICEs] to ask whether there is something they would be missing that we're not. If we were like them rather than ourselves, in other words, would be in error about something, or would we just be different? (p. 294).

The only way for the formal constructivist to avoid this is to reject moral rationalism entirely – both strong and weak versions – accepting whatever costs this may bring. Some will be unmoved by the alleged costs of rejecting moral rationalism. But the costs to constructivism go beyond those that come with rejecting a popular doctrine.

Once moral rationalism has been rejected there is little reason to think that constructivism still poses a serious challenge to realism.¹²² A formal constructivism about practical reason would still conflict with a *comprehensive* ethical realism – i.e., realism about morality, value, and practical reason. But if moral rationalism is false, it turns out that there is little motivation for the realist to be comprehensive. If one is convinced that morality is stance-independent and that moral claims necessarily provide one with reasons for action, one might infer that reasons for action must also be stance-independent. If they weren't, the connection between morality and reasons for action would appear to be contingent. How could there be a necessary connection between things that are essentially distinct? Once moral rationalism is rejected, however, there is little pressure to account for morality and reasons in similar coin. In fact, this alleged difference is at the heart of a skeptical challenge to morality that many realists take seriously.

¹²² Such a concession would appear to concede too much to realists. Recall that one of the motivations for constructivism involves appeal to naturalistic concerns. Constructivists, following Mackie (1977), argue that realism commits one to a queer ontology and epistemology. Arguably, the central premise in Mackie's Queerness argument is that realists are committed to objective prescriptivity. This is supposed to result from taking moral rationalism together with a realist view of objectivity as stance-independence. If the constructivist concedes rationalism, then it is far less clear why she should think that moral properties are *queer*. The charge of queerness would have to apply strictly to the idea that morality is stance-independent. But now the feature of morality that was supposed to make it so different from other stance-independent properties is missing. As the epistemological argument piggybacks on the metaphysical argument, it now turns out that neither objection goes through.

Some realists reject moral rationalism because they take morality and reasons for action to be so different.¹²³ Although they think that realism presents the best account of morality, they deny that moral claims necessarily provide us with reasons for action. One of their motivations for rejecting the doctrine is that it does not take one popular and historically important type of skeptical challenge seriously enough.¹²⁴ Many think that it is an open question whether performing the morally right action is always or often in an agent's best interests. The worry is that reasons for action are plausibly considered egoistically. That is, what reasons an agent has are to be understood in terms of some function of *her* interests. Morality, by contrast, often requires that an agent sacrifice her interests for the sake of others. Since morality will sometimes require self-sacrifice, it would appear as if there is a conflict between morality and self-interest, or prudence. The moral skeptic challenges the realist to show that the nature of morality and the nature of reasons for action are such that we have reason to be moral.

Far from undermining moral realism, a constructivist account of reasons might figure in a realist's response to the skeptic.¹²⁵ In this context, it is not clear that the kind of objections I have been pressing against constructivism turn out to matter

¹²³ See, e.g., Brink (1993).

¹²⁴ For example, this kind of sceptical challenge is voiced by Glaucon and Adeimantus in Book ii of Plato's *Republic*.

¹²⁵ Let us call *rational egoism* the view that one only has a reason if the action contributes to *her* interests. Since claims about reasons, on this view, make essential reference to the agent, we might also refer to this view as an *agent-relative* account of reasons for action. Now constructivism about practical reason also represents an agent-relative account of reasons for action. On this view, an agent's reasons are understood as a function of *her* judgments or stances. So far, this says nothing about an agent's interests. But, as an agent-relative view, there is no reason why such an account could not be made to conform to rational egoism and the terms of the challenge I have presented. In order to square egoism with constructivism, all one needs to do is take an agent's interests to be a function of *her* judgments or stances.

much. For example, suppose that I am correct and it turns out that constructivism cannot *morally* justify the torture claim or rule out the possibility that ICEs like Caligula have no reason not to torture. Even if this is all true, constructivism might still put us in a position to show that *we* (or people like us) have reason to be moral. This would still arguably mark an advance over the view held by a skeptic like Thrasymachus who thinks morality is only for the deluded and the weak – those who cannot see what their true reasons are or who can but lack the courage or resources to act on them. If a constructivism achieves this much, a realist might appeal to a constructivist account of reasons in her response to the skeptic.

A constructivist might avoid the kind of challenge I have presented if she restricts her view to practical reason. But in so doing she incurs new costs. Not only must she absorb the cost of rejecting moral rationalism; she also undermines her own claim to challenge realism. By rejecting moral rationalism, she changes the playing field. In this new context, realism and constructivism might better be seen as complimentary, rather than as opposing, views.

3.8 Conclusion

Current consensus has it that formal versions of constructivism offer the only coherent challenge to a stance-independent realism. But, as I have argued, formal views are the most vulnerable to challenges from within ethical theory. Insofar as they are formal, they lack the resources to justify our considered judgments. The torture claim may stand in as a stock example. But it is not unique. Arguably there are many considered

judgments that take a similar form. That is, we are committed to thinking that many different kinds of acts would be wrong (or permissible, or right, or just, or brave, or cruel, or wicked, etc.) *even if no one ever took them to be such*. Formal views fail to justify these considered judgments *because* they reject (3).

Realism, by contrast, commits one to (3) and would, consequently, appear to do a better job of accommodating these considered judgments. Although they might not guarantee that any particular judgment is correct, there is no principled obstacle to justification for the realist. The realist's success together with the constructivist's failure counts as a comparative advantage for realism. According to the importance that constructivists themselves assign to considered judgments in their theory building, this deficiency significantly undermines the attractiveness of formal versions of constructivism. This result should prompt the constructivist to revisit the plausibility of substantive views.

In Chapters 4 and 5, we will see two different arguments that constructivists have wielded against the realist's condition (4). These would appear to support some version of a substantive constructivism. The first argument concerns moral rationalism and the force of reasons; the second concerns the significance of autonomy to our thinking about moral agency and action. Both arguments are supposed to address the *scope* of condition (3). In other words: what reason do we have to accept the realist's condition (4)?

Realists might disagree with constructivists about particular cases. As I argued in Chapter 1, condition (4) provides a way of explaining this disagreement. But what are the grounds for thinking that one side or the other wins out in these disputes about

ethical objectivity? In other words, why should we accept (4) – viz., the claim that the objectivity characterized in condition (3) extends to *all* moral principles?

Alternatively, are there any countervailing reasons for assigning stance-dependent ethical truths a wide scope in our theory (especially to the extent that Scanlon or Korsgaard suggest)? In the next two chapters, I lay out specific arguments that constructivists have presented in response to these questions. If they are sound, then there is good reason to reject (4) – *even if* (3) correctly characterizes some part of the ethical.

Some constructivists, like Korsgaard, think that this claim follows if we accept moral rationalism, an independently plausible first-order view. They argue that there are certain features of practical reason that resist a realist characterization. That is, these features provide good reason to think that the truth of most reasons judgments is stance-dependent. But if moral rationalism is true, the result is that there is good reason to think that the truth of most moral judgments is also stance-dependent. I present and evaluate this argument in Chapter Four.

In addition to moral rationalism, constructivists also appeal to a distinct, yet related, consideration about the role of autonomy in our thinking about morality and practical reason. They claim that if moral truth were stance-independent, as the realist maintains, we would lack the appropriate control over our actions and who we are. On their view, such a result would leave very little that is recognizable of our conception of morality and moral agency. Autonomy demands that we have control over the standards that govern our practical lives and whether our judgments about them count as true. The constructivist concludes from this that the wider the scope one assigns

stance-dependent ethical truth, the better this is for our autonomy. I present and evaluate this argument in Chapter Five.

CHAPTER 4

NORMATIVE ALIENATION AND THE MORALITY OF COMMON SENSE

4.1 Introduction

Morality can be difficult. It often seems to require that we sacrifice our own good for another's. Such sacrifices may prove costly, perhaps requiring us, on occasion, to lay down our own lives. One of the main challenges for moral theory is to explain why we have reason to be moral, even when it is difficult or contrary to our own interest, and why these reasons should matter to us. In other words, the success of a moral theory will in part depend on its ability to accommodate the following three claims.¹²⁶

- (i) *Commonsense Morality*: Moral requirements include universally applicable other-regarding duties (e.g., duties of forbearance, cooperation, and aid).
- (ii) *Moral Rationalism*: Moral requirements necessarily provide agents with *some* reason for action.
- (iii) *Rational Force*: A conception of practical reason should be able to explain why these reasons matter.

¹²⁶ Cf., Brink (1992).

Although (i)-(iii) are each controversial, I will show that each also enjoys a good deal of intuitive support and that a moral theory that can accommodate all three is preferable, other things being equal.

Christine Korsgaard argues that together these three claims provide her own brand of constructivism a significant comparative advantage over moral realism. In short: they figure in an argument for the rejection of the realist's condition (4).¹²⁷ If Korsgaard is right, all moral truth is stance-independent – including the truth of moral principles.

Very roughly, Korsgaard's own view is that (a) commonsense moral requirements are requirements of practical reason, (b) the requirements of practical reason are constituted by procedurally sound deliberation, and (c) deliberation counts as procedurally sound just in case it conforms to the requirements of agency itself.

The first claim is supposed to explain (ii) or why the connection between moral requirements and reasons for action is a *necessary* one. It says that the moral truths are just a subset of the truths about practical reason. This means that anytime a moral requirement applies to an agent, this agent will also have at least some reason to perform the action that morality requires of her. The second claim makes Korsgaard's view a species of *constructivism*. It says that the truths of practical reason (which according to the first claim include moral truths, as well) are in some sense *constructed* by us in deliberation. It is also supposed to explain (iii) or why we should care about rational demands. On this view, rational demands (and, hence, moral demands) command our assent because they are themselves the outcomes of our own

¹²⁷ See Chapter 1, section 5.

engaged deliberation – or, at least, what our deliberation ought to be, so long as we are deliberating well. The third claim about agency is supposed to explain (i) or why our actual moral requirements conform to common sense. How and whether this last claim in fact works is controversial.¹²⁸ But in order to simplify things here, I will assume that this is true. I will also assume that the first claim works to satisfy (ii). My focus, rather, will be how the second claim secures (iii) and, to the extent it does, whether it in fact does a better job of this than moral realism, as Korsgaard contends.

A robust moral realist, by contrast, is allegedly only able to accommodate two of the three claims at best. As I stated it in Chapter 1, robust moral realism involves an especially strong claim about *objectivity*. On this view, *all* moral principles are supposed to be objective in the sense that the matter of their truth is independent of the intentional *stances* (e.g., beliefs, desires, commitments) we bear toward morality. That is, whether a general moral statement is true or not (e.g., stealing is wrong) is independent of whether we believe it, or desire it, or would agree to it, etc. In other words, we don't make these moral principles true just by believing that they are true or desiring or wishing that they were, or agreeing to act in accordance with the actions they propose. It is the notion of moral objectivity expressed in condition (4) that is supposed to cause a problem for realism, one that allegedly gives us reason to reject robust versions of the view.

Whereas we might assume that any version of realism will account for (i), Korsgaard argues that the realist's commitment to (what I have been calling) condition

¹²⁸ The difficulty that constructivists have in establishing this claim is illustrated in Chapter 3. Although the topic there was formal constructivism, it would appear as if similar problems may arise for defenders of substantive views.

(4) is incompatible with both (ii) *and* (iii). That is, she thinks a realist may account for either Moral Rationalism or Rational Force, but not both; she has to choose. If this is correct and constructivism is able to accommodate all three claims, this would count as a significant comparative advantage for Korsgaard's view. For constructivism is supposed to secure a sense of moral objectivity *without* (4).

For the sake of argument, I will assume that the considerations expressed by (i)-(iii) *are* plausible constraints on a moral theory. My approach in this paper will be to show that even if we accept this much, we have no good reason to accept Korsgaard's constructivism. This is because a realist can exploit the same general strategy of accommodation that Korsgaard's constructivist employs. If I am right, (i)-(iii) turn out to be at least as compatible with a robustly realist, stance-independent account of moral truth as they are with constructivism.

In section 2, I will try to clarify and motivate each of our three initial claims. In section 3, I will then present Korsgaard's argument for the claim that realism is incompatible with these considerations and show exactly how constructivism is supposed to avoid a similar conclusion. In sections 4 and 5, I will return to the topic of Rational Force. A more careful review of the things that Korsgaard has said about this condition will reveal some problems for her own view. Specifically, Korsgaard sometimes describes the requirement expressed by Rational Force in terms of motivation; other times, however, she describes it in terms of something that could serve to block a potential regress of practical justification. As I will argue, however, Korsgaard's own view appears inadequate to satisfy the condition of Rational Force when it is understood in either of these ways. This failure prompts me to ask whether

there might not be some third way of construing Rational Force, one that Korsgaard's view in fact satisfies and yet which might still serve to contrast her view against (some versions of) realism. In section 6, I claim to have identified just such an account – one that appears to enjoy some independent plausibility. However, as I go on to show in section 7, even this characterization of Rational Force would only appear to give constructivism an advantage over *some* versions of moral realism, not all. If this is right, then Korsgaard is mistaken. These three considerations do not provide constructivism with a comparative advantage over realism per se. Even if they might show that constructivism is superior to some robustly realist views, constructivism does not turn out to be superior to every robust version of realism.

4.2 The Explanatory Challenge

In theorizing about morality, our aim in part is to systematize our deeply held moral judgments at various levels of generality.¹²⁹ This includes both considered judgments about particular cases as well as the theoretical judgments that might help to explain or systematize these. As I will soon explain, each of (i)-(iii) is initially quite plausible and arguably represents a commitment that we should be concerned to accommodate in our moral theories. Hence, a theory with more explanatory power than another with respect to these commitments is better, other things being equal.

¹²⁹ Again, the method of theory construction I have in mind here is *reflective equilibrium*. Every major defender of constructivism accepts this method (see Chapter 3). Hence, it is fair to assume it here. Again, I say the aim is only *in part* to systematize considered moral convictions since one might also be concerned to bring one's moral theory into line with one's considered theoretical commitments about other topics – e.g., one's commitments in metaphysics, epistemology, or the philosophy of language.

Commonsense Morality includes all of those other-regarding duties that potentially make being moral so difficult. According to Common Sense, you have a duty, amongst others, *not* to harm or take advantage of others, as well as a duty to keep your promises, give others what is owed to them, and aid others when the threat to them is great and the cost to you insignificant. Now, it may turn out on occasion that certain features of one's particular situation are such that the obligation that usually attaches to these principles of action is outweighed or cancelled. But, in general, I take it as a platitude of our everyday moral thinking that there are moral requirements that *are* other-regarding in something like this way and that the actions they require represent potential sources of sacrifice.

A quick qualification: One need not think that Common Sense is *infallible*. We may discover, after much reflection and inquiry into the relevant sort of facts, that certain commonly held beliefs about what our duties are turn out to be mistaken – i.e., that certain so-called “duties” are not in fact moral requirements. For example, we might not have a duty to keep our promises per se; rather, it might turn out that we only have a more general duty to show respect for ourselves, and others – and this usually involves the keeping of promises, but it needn't in every case. Or it might turn out that there is no demanding duty of aid to others with whom we share no significant bonds or relations, even if the threat to these others is significant and the cost to us of giving aid insignificant.¹³⁰

However, it would also appear as if there is a limit to how much is revisable.

Although Common Sense is compatible with *some* revision, it does not seem

¹³⁰ This is a view that might appear especially plausible when the question is whether, or the extent to which, we have a demanding duty of aid to the global poor or distant future generations.

compatible with the kind of wholesale revisionism that many see in certain kinds of challenges to morality: e.g., ethical egoism, or say a view that committed us to the rejection of justice as a general moral requirement.¹³¹ There might be compelling arguments for these views. But, nevertheless, they would still lack one feature that Common Sense has in abundance: viz., intuitive appeal. What makes Common Sense plausible is simply that the kinds of duties it enumerates are just the kind of moral requirements that people would commonly accept after minimal, sober reflection on morality.

As we know from Chapter 3, *moral rationalism* is the view that moral requirements necessarily provide the agents to whom they apply with some reason for action.¹³² It is a consequence of this view that if an action is morally wrong, we have at least *some* reason not to perform or promote it.¹³³ Moral rationalists take this to be a

¹³¹ According to *ethical egoism*, morally right action for any agent is one that would produce certain good consequences for the agent performing the action. One worry is that this sort of view is incoherent in some way. However, another – and the one that concerns us in this context – is that egoism is incompatible with Commonsense. Whatever duties we have to others, on this view, are always derivative. That is, they only count as duties to the extent that they contribute to the conferral of good consequences for the agent. This view arguably does not capture the importance of *impartiality* in our moral thinking. Commonsense, by contrast, states that these duties are impartial – i.e., more than merely derivative – even if they could be justified by reference to the agent’s own good or interests, as well. One might also worry that utilitarianism or other versions of consequentialism pose a similar threat to Commonsense. However, by contrast with egoism, the threat is less clear. Although defenders of these views are often open to significant revision of our moral commitments – consequent on the what the outcomes turn out to be, utilitarians (or consequentialists) are concerned to accommodate impartiality, as well as a good deal of Commonsense. Perhaps an exception to this last claim would be the *scalar consequentialism* advanced most prominently by Alastair Norcross (2006) – see also Howard-Snyder, Norcross (1993). According to this view, we ought to jettison the deontic vocabulary that informs our Commonsense talk of duties and restrict ourselves to the notion of relative value – i.e., the idea of an action producing outcomes that are *better or worse than* those produced by another available action. Such a view would appear to conflict with Commonsense.

¹³² See, e.g., Shafer-Landau (2003) and Finlay, Schroeder (2008). Brink (1989) refers to this view as *internalism about reasons* (see, pp. 39ff). I opt for *moral rationalism* to avoid confusion with a different thesis that bears the same name: Bernard Williams’ (1979) view that what reasons an agent has depend on the elements of her motivational set.

¹³³ Sometimes it is claimed that morality necessarily provides overriding reasons for action. However, this strong claim is difficult to accommodate theoretically. See Brink (1997). It is also less clearly a deep feature of our ordinary moral thinking and practice. It looks as if there can be genuine conflicts

necessary truth.¹³⁴ Anti-rationalists deny this. They think that what reasons we have and what moral truths obtain are independent questions, overlapping only contingently, if ever.¹³⁵

Moral rationalism is not uncontroversial but it commands a good deal of intuitive support. Consider the judgment that torturing suspected terrorists is morally wrong. When we deem this kind of behavior morally wrong, we imply that people who torture suspected terrorists are in violation of a standard of appropriate conduct. But, if moral rationalism were false, these standards would not *themselves* supply reasons for action. In other words, torturing suspected terrorists may be immoral, or morally unjustified, or inappropriate, but nevertheless there may turn out to be no reason not to perform this kind of action. And this seems confused. It seems mistaken to claim that the government, for example, ought not to torture suspected terrorists – and is morally wrong to do so – while also claiming that there is no reason at all for the government not to torture. What’s more, it might seem unfair to criticize someone who attends to all of her reasons for action, gives them their appropriate weight in

between moral reasons and reasons of other kinds – e.g., reasons that relate to art, or the particular allegiances one has to one’s family or country – and that the moral ones are not always guaranteed to win out. This is in part why some moral decisions prove so difficult. Korsgaard recognizes this possibility herself. See Korsgaard (1996): 125-6. For this reason, I restrict myself to a weaker formulation of moral rationalism, one that states that moral claims necessarily provide *some* reason for action, even if these reasons ultimately fail to win out in deliberation.

¹³⁴ Sometimes moral rationalism is also understood as an *a priori* conceptual claim. Brink (1989) presents the view this way. It says that the nature of our justification for believing statements like the one above is *a priori*. The necessary connection between morality and rational authority is something that we come to understand through reflection on the nature of our moral concepts, not by discovering facts out in the world or through experience. This is because it represents a *conceptual* truth. Someone who judges that it would be wrong for an agent *S* to ϕ but denies that *S* has any reason to avoid ϕ -ing is not guilty of possessing a mistaken moral belief. That is, she does not merely fail to apply the concept of *wrongness* correctly. Rather, she fundamentally fails to grasp what *wrongness* is; she does not possess the concept. However, these further specifications turn out not to matter for the debate I present here. As Korsgaard presents the challenge, it is the necessity claim that realists fail to explain. For simplicity, I restrict my focus to the necessity claim.

¹³⁵ Foot (1972); Brink (1989)

deliberation, and acts accordingly. If the government acts on all of the reasons available to it and has no reason not to torture, it seems morally inappropriate to criticize it. These aspects of our everyday moral thinking appear to support the truth of moral rationalism.¹³⁶ However, moral rationalism does not itself explain why we should care about conforming our actions to rational demands. Rather, this is something that is addressed by the next thesis.

Rational Force states that a good ethical theory must provide some explanation for thinking that we should care about conforming our actions to rational demands that does not simply involve supplying another *reason*. Reasons for action, unlike brute desires and impulses, have a special authority that is supposed to matter to our deliberations. They *ought to* grip us and command our assent. But if our reasons bore no particular relation to the conditions of our actual deliberations or the kinds of considerations that in fact matter to us, it would be difficult to see why these rational demands should *in fact* grip us or command our assent. What we want is an account of what a reason *is* and why we have them that explains why rational demands matter – why we (should) *care*.

Now, as we are going to see, it is very difficult to say exactly what this all means or how this explanatory condition constrains theory. Specifically, what do we mean by *care* here? And, once we get clear on this, how strong is this requirement? Must an account explain why we will *necessarily* care in the relevant sense? Or will

¹³⁶ In short, if moral rationalism is true, then one *cannot* be held in violation of a moral standard unless one has at least some reason not to perform or promote a particular action. There seems to be some confusion involved in supposing that one who violates a moral standard need not act contrary to any reason she has. Conversely, if someone has no reason to refrain from an action, it seems unfair to blame her for performing it. See Shafer-Landau (2003), 192-3.

merely *some* account of the connection between reasons and an agent's concerns and interests satisfy this condition?

Let this statement of the condition serve as something like a *first pass* (one that we will revisit as the discussion unfolds). Even if it still needs refinement, as I think it does, it is clearly in the neighborhood of a genuine and legitimate ethical concern that someone might raise. If someone tells you that you morally ought to do something, I think it is reasonable to ask why we should care? The extent to which a view of morality or reasons for action can address this kind of question will thus matter to its plausibility on the whole.

So far, I have been concerned to explain and motivate (i)-(iii). From this brief sketch of the theses, I hope it is clear why many philosophers have found them plausible. Although a view that can accommodate each of these three theses is preferable, other things being equal, it turns out that it is very difficult to accommodate all three.

4.3 The Argument from Moral Reasons

As I've already mentioned, Korsgaard thinks that moral realism is incompatible with all three of these considerations and that, at best, it can accommodate only two of them. The problem with realism, on her view, is exactly that feature that distinguishes it from her own brand of constructivism: viz., the realist's commitment to the stance-independence of *all* moral principles. She argues that if we take this realist

commitment *together* with Moral Rationalism, the pair turns out to be incompatible with an account of Rational Force.

Although it is widely agreed that Korsgaard presents this argument in several of her works, it is difficult to find a concise or precise statement it.¹³⁷ What follows here is my own reconstruction of what I take her argument to be, one that I think represents both a fair and compelling interpretation of at least one of the many things that she finds troubling about moral realism. This presentation will also prove useful insofar as it should help us to identify and isolate the claims that are doing the heavy lifting and, consequently, put us in a good position to see what the most plausible route for a realist response will be.

Argument from Moral Reasons

1. All true moral principles are essentially stance-independent. [Robust Moral Realism]
2. Moral requirements necessarily provide agents with *some* reason for action. [Moral Rationalism]
3. If the nature of morality were essentially distinct from the nature of practical reason, there could be no necessary connection between the two. [Hume's Dictum]
4. So, the nature of morality and the nature of practical reason are not essentially distinct. [2, 3]
5. So, the truths of practical reason are also ultimately stance-independent. [1, 4]

¹³⁷ For the fullest treatment see Korsgaard (1996).

6. If the truths of practical reason are stance-independent, they bear no immediate or necessary connection to an agent's deeply held concerns. [Korsgaard's Assumption]
7. So, realism cannot account for the Rational Force of moral reasons. [5, 6]

If the conclusion of this argument is correct, moral realism fails to satisfy Rational Force, an important constraint on our moral theorizing. What's more, if Korsgaard is justified in claiming that her own view *can* satisfy this constraint, it would appear as if constructivism bears an important comparative advantage over realism. I should add that I think that this argument is a *valid* one – i.e., if the premises are all true, we have no choice but to accept the conclusion. This means that if we are going to reject the conclusion we must show that at least one of the premises is *false*.

As I've presented it, the argument includes four independent assumptions – viz., 1, 2, 3, and 6. The rest of the premises – viz., 4, 5 – are supposed follow from them. Hence, if we are going to show that one of the argument's premises is false, this is ultimately going to come down to one of these assumptions being mistaken. In what follows, I'd like to make a few comments about each of these and provide some reasons that might be thought to support them before moving on to discuss my own strategy of response.

The first assumption, in premise 1, is just supposed to follow from a commitment to moral realism.¹³⁸ It represents a constitutive feature of the view that Korsgaard intends to target. As I mentioned at the beginning of the discussion,

¹³⁸ See Chapter 1, sections 5 and 6.

philosophers disagree about whether to include a stance-independence condition as part of realism, as well as whether it should take this particular scope. But, without this assumption, it does not appear as if Korsgaard has an opponent in this debate. It should also be noted that there are self-proclaimed realists (e.g., Russ Shafer-Landau, Derek Parfit, William Fitzpatrick) who insist on something like the view expressed in premise 1. What's more, these realists think that it is *because* of this commitment to 1 that they are able to lay claim to Commonsense Morality. They think that *because* morality is independent of us and our minds in this particular way, there is no special worry about how moral requirements could be other-regarding in the way that Common Sense suggests.

Korsgaard, however, thinks that this commitment of realism is what leads to the argument's troubling conclusion and, consequently, what counts in favor of her own constructivism. Recall that constructivism entails that at least *some* moral principles are stance-dependently true. Importantly, Korsgaard thinks that her own view avoids the argument's conclusion because it entails the rejection of premise 1.

The second assumption, in premise 2, is one that we have already had some occasion to discuss. It is just Moral Rationalism. As we have seen, this thesis would appear to be well supported by our everyday moral thinking and practice. Of course, one might worry that it might not be possible to accommodate *all* of the kinds of considerations that appear plausible to us in the everyday. Perhaps one of the main reasons we engage in philosophical reflection about morality is to help us decide how to best revise our moral commitments so that they in fact cohere with each other. If

this is right, then we might imagine a realist deciding to reject premise 2. In fact, some do.

Some realists (including, e.g., David Brink) reject moral rationalism because they take morality and reasons for action to be so different.¹³⁹ Although they think that realism presents the best account of morality, they deny that moral claims *necessarily* provide us with reasons for action. One of their motivations for rejecting the doctrine is that it does not take one popular and historically important type of skeptical challenge seriously enough.¹⁴⁰

Specifically, they think that it is an open question whether performing the morally right action is always or often in an agent's best interests. The worry is that reasons for action are plausibly considered egoistically. That is, what reasons an agent has are to be understood in terms of some function of *her* interests. Morality, by contrast, often requires that an agent sacrifice her interests for the sake of others. Since morality will sometimes require self-sacrifice, it would appear as if there is a conflict between morality and self-interest, or prudence. The moral skeptic challenges the realist to show that the nature of morality and the nature of reasons for action are such that we have reason to be moral. A realist might respond to this challenge with an account of why, in the actual world, it works out that agents usually have reason to be moral. But because she thinks that this kind of skeptical challenge is a genuine one, she rejects the view that the connection between morality and practical reason is

¹³⁹ See, e.g., Brink (1993).

¹⁴⁰ For example, this kind of skeptical challenge is voiced by Glaucon and Adeimantus in Book ii of Plato's *Republic*.

necessary. In other words, this kind of realist accepts Common Sense and Rational Force but rejects Moral Rationalism.

Although these kinds of considerations might give a realist good reason for rejecting premise 2, we will see that there are some strong considerations on the other side. What's more, Korsgaard herself accepts this premise. Her commitment in itself doesn't provide any special reason for thinking that premise 2 is true. But if we can show that it is possible to get everything she claims to care about without rejecting realism, we will have given her an especially strong rebuttal. In other words, we will have shown that the argument fails on her very own terms.

The third assumption, in premise 3, is a metaphysical claim. It finds support in a principle that David Hume is supposed to have maintained in the *Treatise*.¹⁴¹ Because of this connection, it is sometimes referred to as *Hume's Dictum*. In its more general form, very roughly, it says that *there are no metaphysically necessary connections between distinct types of entities*.

I take it that the basic idea here is that for a metaphysically necessary connection to obtain between two things, there has to be some kind of explanatory basis for that connection. But if the things belong to distinct types of entities, there is no such explanatory basis. So, there are no metaphysically necessary connections between such distinct types of things.

Now, this is a principle that has been widely accepted. But it turns out that very little has been said in its support.¹⁴² Some philosophers have claimed that it is *directly*

¹⁴¹ Hume (2000), Book I, Part III, § VI.

¹⁴² For discussion of some of the standard justifications that have been given, as well as a novel approach, see Jessica Wilson (2010).

justified – either in virtue of following analytically from the concepts involved or as simply “...expressing intuitions we have no good reason to question.”¹⁴³ This kind of justification is not going to satisfy anyone who doesn’t already intuitively appreciate the force of the principle. But as there is also nothing particularly problematic about this assumption for realism per se, I will ask that we simply take it for granted as a bit common ground. Korsgaard herself does not explicitly appeal to it, but she also doesn’t say anything that would count against it. In any case, I think we need to supply it in order to make the argument go through. This leaves us with one last assumption to target.

The fourth assumption, premise 6, is what I have called *Korsgaard’s Assumption*. I take it that this conditional statement is supposed to present a necessary condition for accommodating Rational Force, one that realism *allegedly* fails to satisfy. But why should we accept this condition?

Our earlier discussion of Rational Force *does* appear to support the claim that a successful account of practical reason will be one that shows how reasons are immediately or necessarily connected to an agent’s deeply held concerns and interests. This much is uncontroversial. But what is not clear is why this requires that the truths of practical reason *not* be stance-independent. In order to justify *this* claim we must take a closer look at Rational Force. If what we have said on behalf of the other assumptions has been convincing, then premise 6 would appear to be the only vulnerable assumption in the argument. Hence, if we want to reject the argument’s conclusion, we had better be able to find a problem with Korsgaard’s Assumption.

¹⁴³ Wilson (2010), 596.

In what follows, I will start off by canvassing some things that Korsgaard has said about Rational Force that might be thought to support this assumption. However, as I will argue, the refinements to Rational Force that they suggest are also ones that Korsgaard's own view (viz., constructivism) *itself* fails to satisfy. This result would, of course, also undermine her claim that constructivism enjoys a comparative advantage over realism. This failure prompts me to search for some further refinement of Rational Force that might do the trick. Although I think that I have found just such a view – or at least one that would appear to give constructivism a comparative advantage over some versions of realism, it is not clear in the end that even this version will serve to support Korsgaard's Assumption. This is because there would still appear to be other versions of realism that can similarly satisfy this characterization of Rational Force.

4.4 Rational Force: *A Second Pass*

Korsgaard talks about the requirement of Rational Force in different ways. On the one hand, she sometimes talks about it as the requirement that an account of practical reason explain the apparent connection between an agent's having a reason and that agent's motivation to act accordingly. On the other hand, she also talks about it as a requirement that the view be able to block a certain kind of skeptical challenge – one in which the skeptic responds by asking “But why should I care about reasons like *that?*” But, as I will argue, Korsgaard's own view fails to satisfy Rational Force when it is construed in either of these ways.

Let's first consider the construal of Rational Force in terms of some kind of motivational requirement. Here the basic idea is that when we ask why we should care about conforming our actions to the requirements of practical reason, the relevant sense of *care* is one of motivation. In other words, on this view, we might construe Rational Force in the following way:

- (iv) A conception of practical reason should be able to explain why an agent will *necessarily* be motivated to act according to the reasons she has.

A few comments here: First, it is important that the connection between reasons and motivation be a *necessary* one. Anything less would make the condition too weak to be interesting. If the condition merely required an explanation for why an agent *might be* motivated to act according to the reasons she has, there would be no great challenge for realism (or any other view of morality). People might be motivated by all kinds of considerations. There is no reason to think that the moral should be any different. Second, however, (iv) does not say that the account must explain why an agent will necessarily be *moved to action* on the basis of her having a particular reason; rather, it should be read to say that she will have *some* motivation to perform the actions her reasons prescribe. This qualification is important, since – on at least some views – an agent might have a great many *pro tanto reasons* (i.e., reasons that count to this or that extent, but ones that might equally fall short of *all-things-considered reasons for action*). Of the many pro tanto reasons that apply to an agent in any particular

situation, some will prescribe practically inconsistent courses of action. This would of course be problematic if each necessarily moved her all the way to action. But so long as we take (iv) to say that the agent will only be necessarily motivated to some extent, no such problem arises.

This understanding of Rational Force helps support the conclusion of our initial argument against realism. That is, it appears to lend some support to Korsgaard's Assumption. If realism commits us to the view that the truths of practical reason are independent of our intentional stances, it also commits us to the view that these truths will be independent of an agent's motivational stances. This is because motivational stances (e.g., desires, plans, commitments...) are a kind of intentional stance. So what an agent has reason to do and what she will be motivated to do may come apart. In fact, it is at least possible, on this view, for an agent to have a reason but be in no way motivated to perform the action it prescribes. If one, however, is committed to there being a *necessary* connection between reasons and motivation, the realist's view of practical reason will turn out to be a disappointment. In short, the realist cannot satisfy (iv).

In this respect, one might think that Korsgaard's view bears an advantage over realism. But a quick inspection reveals that her constructivism can do no better with respect to motivation – even though it does not commit one to the allegedly problematic first premise of the argument. Whereas a rejection of the first premise is sufficient to block the conclusion to *that* argument, it turns out that there is another argument that starts with (i) and (ii) and forces us to reject (iv).

Here's the argument. Suppose we accept Common Sense and Moral Rationalism. That is, we accept that moral requirements include universally applicable, other regarding duties *and* that these requirements necessarily provide the agents they apply to with some reason for action. From this it follows that some moral reasons are universally applicable and other regarding. But if our reasons (including these moral reasons) depended on agent's desires or contingent motivational set, a particular agent might not turn out to have such reasons. But we just said that we do have such reasons (per stipulation). So, our reasons (including our moral ones) do not depend on an agent's desires or contingent motivational set. So, it would appear as if someone who accepts (i) and (ii) – like Korsgaard – is unable to also coherently accept (iv) or Rational Force when it is construed in terms of motivation. We might say that (i) and (ii) commit one to the view that moral reasons are *categorical* – in the sense that they are not contingent on an agent's desires or motives.

Of course, to the extent that one finds (iv) a plausible way of cashing out the requirements of Rational Force, one might conclude that this gives us some reason to reconsider our commitments to (i) or (ii). Earlier, I mentioned Brink and Foot as those who reject (ii). But there are also those who go the other way and reject (i).

For example, a moral relativist – like Gilbert Harman – might accept (ii) and the motivational construal of Rational Force in (iv) but reject (i).¹⁴⁴ A relativist might try to account for the force of practical rationality by reducing claims about reasons to claims about an agent's motivational states. On this view, we care about the demands of practical reason because these demands are just a function of our caring about

¹⁴⁴ Harman (1975).

certain things in the first place (i.e., our antecedent motivational states). In fact, this reduction of facts about reasons to facts about motivation is what one might think Hume's Dictum requires in order to explain the necessary connection described in (iv). However, since different people start off with different concerns and priorities, the demands of practical reason might differ across agents. In fact, on this view of practical reason, it might turn out that an agent has no reason whatsoever to aid, forbear, or cooperate with others in the way that Commonsense morality dictates. For example, on this view of practical reason, we might assume "...that, for Hitler, there might have been no reason at all not to order the extermination of the Jews."¹⁴⁵ Together with moral rationalism, we would then have to conclude, as Harman in fact does, that "...it sounds odd to say that Hitler should not have ordered the extermination of the Jews, that it was wrong of him to do so."¹⁴⁶ But I think that, for many, this is a bullet that is just too big to bite. What's more, the stronger rebuttal involves showing that Korsgaard fails on her own terms. And, importantly, Korsgaard accepts Commonsense Morality.

Insofar as Korsgaard accepts (i) and (ii), she is also committed to the view that moral reasons are categorical in the sense that they are *not* contingent on an agent's particular desires or motives. But then we have not vindicated her claim to accommodate all three of our considerations; we have, as of yet, no reason to think that her own constructivism betters moral realism. Either her own constructivism fails to meet the challenge she presents or we must understand Rational Force in some other sense.

¹⁴⁵ Harman (1975), 9.

¹⁴⁶ Harman (1975), 7.

4.5 Rational Force: *A Third Pass*

So the motivational construal of Rational Force asked us to provide an account of reasons for action that could explain why someone who has a reason will necessarily be motivated to conform her actions to them. But – as we just saw – it is not just the realist who has a problem with this condition. Rather, any view that commits one to (i) and (ii) will fail to satisfy this kind of motivational requirement – and this importantly includes Korsgaard’s own view.

But there is another way of characterizing Rational Force that we should consider – one that Korsgaard also appeals to at times.

- (v) A conception of practical reason should make any skeptical challenge unintelligible.

This condition states that a successful account of practical reason will be one that explains what a reason is (and why we have them) in such a way that one cannot intelligibly ask why reasons like that should matter. In other words, once one is presented with the account, the question “But why should we care about *that*?” should not have an *open* feel to it; it should not express an intelligible question or one that any person who had a command of the relevant concepts could reasonably ask. If, however, the account may be met with open questions, it fails to satisfy this condition for Rational Force.

This approach represents an attempt to excavate the core truth out of G.E. Moore's famous *Open Question Argument*.¹⁴⁷ Although it is now generally acknowledged that Moore's own version of the argument (one which relies on a controversial synonymy test for property identities) fails in one respect or another, some philosophers – including Korsgaard – nonetheless take Moore to have been onto a deeper objection against certain reductive views of normativity and Rational Force. In fact, Korsgaard herself has said that her argument against realism “is the ultimate extension of [Moore's] open question argument.”¹⁴⁸

Korsgaard argues that any substantive realist account of practical reason fails to meet this condition. Take, for example, a paradigmatic yet admittedly caricatured view of moral reasons – a view we might call *Unrefined Platonism*. On this view, let's say that what moral reasons we have are determined (in some fashion) by the relation our actions bear to certain supersensible entities – *the forms*. Korsgaard thinks that this kind of account of moral reasons does not satisfy the open question test. Here's why.

Suppose I tell you that it's wrong to torture suspected terrorists and that you have reason not to perform actions of this type. You might just respond, “So what? Why should I care about conforming my actions to these reasons?” At this point, our Unrefined Platonist might reply, “Because, if you consult *the forms*, you will see that you have such reasons.” To which you respond, “But why should I care about *that*?” – “Why should I care about what *the forms* say I should do?”

On the one hand, Korsgaard quite plausibly thinks that these last two questions remain open and that, as a result, the view fails to satisfy (v). Not only do they express

¹⁴⁷ Moore (1903).

¹⁴⁸ Korsgaard (2003), 112.

intelligible questions, it is also not clear how any realist could satisfactorily block this line of questioning. In this respect, (v) would also appear to support Korsgaard's Assumption. One might think that a view counts as immediately or necessarily connected with an agent's deeply held concerns and interests *only if* it successfully withstands an application of the open question test. But since no stance-independent view satisfies the test, it follows that no such view counts as satisfying the constraint of Rational Force.

But, on the other hand – and far less plausibly – Korsgaard takes it that her own constructivist account *does* block this kind of skeptical rejoinder. But it is not clear that it can. Again, as with the last candidate, the condition expressed by (v) appears too strong for Korsgaard's own view.

Here, the relevant question is something like the following: “But why should we care about conforming our actions to the results of procedurally sound deliberation?” – “Why should we care about reasons like *that?*”¹⁴⁹ These last two *do* appear to be open – at least in the sense that they appear *intelligible*. Someone who responds to the constructivist's account in this way does not appear to be guilty of any kind of obvious confusion.

But this should not be surprising. The condition expressed by (v) would appear to be one that is very difficult to satisfy. Just think of how easy it is to generate intelligible instances of this question form. In fact, this kind of argument is sometimes

¹⁴⁹ We might grant that the question “But why do we in fact have *reason* to perform actions that we would decide to perform as a result of sound practical deliberation?” is *not* open. But this would not count as a relevant application of the test. The test is intended to support the claim to Rational Force. Some philosophers – notably those non-naturalists I will mention in a moment – would not grant this.

employed to show that *no* substantive account of reasons is plausible and that the very demand for an explanation of Rational Force is confused.

Defenders of non-naturalist versions of realism – e.g., Enoch (2007), Fitzpatrick (2008), Parfit (forthcoming), (2006), (1997), Shafer-Landau (2003) – accept both Commonsense and Rationalism but reject the idea that we can or need to explain the force of practical reason – i.e., they accept (i) and (ii) but reject (iii). They do this because they think (iii) would involve something like (v), and as we’ve seen it is not clear that any substantive view of practical reason will be able to satisfy the open question test. These realists think that moral requirements necessarily provide us with reasons for action – and that these reasons matter. But they argue that any attempt *to explain* the force (i.e. normative significance) of these reasons ultimately threatens to rob practical reason of the very force one sets out to account for.¹⁵⁰ Although they think that we may put ourselves in a position to appreciate the rational force of moral requirements by attending more carefully to certain considerations or working through moral arguments, they do not think that we can provide any satisfactory explanation for why they have this force.¹⁵¹ Rather, on their view, moral reasons matter because the kinds of demands they present have *intrinsic* (and unexplainable) rational force.

¹⁵⁰ It is common for these realists to support this claim by appeal to something like G.E. Moore’s *Open Question Argument* – see Fitzpatrick (2008), 175-8; or Parfit (forthcoming), esp. Chapter 26, “The Fact Stating Objection”. Although these philosophers generally acknowledge that Moore’s exact version of the argument fails in one respect or another, they nonetheless take Moore to have been onto a deeper objection against reductive views of normativity. For criticism of Moore’s original argument see Prior (1949), Brink (1989), (2001), and Sturgeon (2003). Attempts to excavate a deeper insight from Moore’s failed version of the argument have also been made by non-realists, like Gibbard (2003) and Korsgaard (1996).

¹⁵¹ Parfit (2006):

Non-reductive realists, as I have conceded, do not give helpful answers to these questions. According to them, we can explain some normative concepts, but only by appealing to others. ... But normative concepts cannot be explained in non-normative terms. Nor can we say much

Perhaps these non-naturalists are correct and *no* substantive view of practical reason – including Korsgaard’s constructivism – satisfies (v). Nevertheless, this might not be a problem that Korsgaard should be concerned with.¹⁵² We might worry that the kind of open question test that (v) involves is *too* strong and that this provides some reason for rejecting it.¹⁵³ As I established in section 2, Rational Force expresses a plausible constraint on our ethical theorizing. But the open question test shows that those who accept this constraint are mistaken or somehow confused. To the extent that we find this conclusion surprising, we might just think that we haven’t yet settled on the right understanding of Rational Force.

One conclusion to draw from this discussion and the one in the previous section would be that Korsgaard cannot secure Rational Force. Another – and one I will soon develop – is that that we should understand Rational Force in yet another way. After discussing what this might amount to and how Korsgaard’s own view might satisfy it, I will entertain the further question of whether there might be some version of realism that can similarly satisfy it.

to explain how we understand these concepts, or how we recognize normative truths. ... We can ask what normative concepts, such as *ought* and *reason* mean. But there are no answers to these questions that are both interesting and true. (pp. 330-1).

¹⁵² It is for this reason that I will forgo a detailed reconstruction and evaluation of Korsgaard’s response.

¹⁵³ Surely, the mere fact that a candidate explanation of practical reason is subject to doubt does not militate against our eventual acceptance of this explanation on holistic grounds. I will not attempt to settle this complicated issue here.

4.6 Rational Force: *A Last Pass*

Although I think that Korsgaard fails to accommodate all three of our initial considerations when Rational Force is understood in terms of motivation (as (iv) characterizes it) or as a response to a potentially open question (as (v) characterizes it), I think that a closer look at her own account helps to shed light on what other kind of requirement she could have in mind. What's more, I think that such a requirement enjoys some independent support.

- (vi) A conception of practical reason should explain the nature of reasons in a way that avoids the threat of *normative alienation*.

In order to understand this constraint and how it contributes to an account of Rational Force, we need to address the following questions: What is normative alienation? Why is it something to be avoided? How might an account avoid it?

Before we ask what *normative* alienation amounts to, we might start out asking what we mean by *alienation* in general. The concept I am concerned with here is complex but not ultimately technical. Like many of our concepts, *alienation* might end up capturing a variety of related but ultimately distinct ideas. I won't presume here that I will be able to give it any precise formulation. Following Peter Railton, we might just say that "...at a perfectly general level alienation can be captured very roughly as a kind of estrangement, distancing, or separateness (not necessarily

consciously attended to) resulting in some sort of loss (not necessarily consciously noticed).”¹⁵⁴ Common examples might include the following kinds of cases:

Relocation

You move to a new country and find, after initial attempts to assimilate and settle in, that you are just not at home with the customs and habits of your new community. You feel like an “alien”... like nobody understands you. You feel as if you lack the kind of emotional support you enjoyed back home. Consequently, you feel worse off.

Intemperance

You’ve had a bad day at work. You feel exhausted, underappreciated, and -- as a result -- quite irritable. As you stand waiting in the checkout line at the grocery store, someone cuts the line in front of you. But this is just too much. You let loose a torrent of angry epithets, reducing the queue cutter – who is caught quite unaware – to a muttering puddle. But once you’ve finished you too stand there equally surprised and shocked. “How could I have done that?” you ask. You think, “That’s not who *I* am.” You feel as if you’ve compromised your integrity by letting your anger get the better of you. You can’t identify with this kind of person. We might say that you feel *alienated* from yourself.

Normative alienation shares with these cases the idea of loss as a result of estrangement, distancing, or separateness. In this case, however, the cause is not a

¹⁵⁴ Railton (1984), 134.

relocation or an outburst of anger, but a set of alien demands.¹⁵⁵ And the related loss does not involve community or integrity; rather, it is the loss of what I have been calling *Rational Force*. More specifically, one might worry that a view that offered no explanation of the connection between an agent's reasons for action and the agent's concerns or interests – i.e., the kinds of considerations that actually appear to figure in an agent's deliberation – would appear to alienate the agent from the reasons that are supposed to obtain for her.¹⁵⁶

I take it that a characterization of Rational Force in terms of normative alienation, as I've sketched it here, enjoys a couple of advantages. First, normative alienation is something that is intuitively very bad and to be avoided. Although there might be cases where a bit of alienation could turn out to be a good thing (e.g., a case where a bit of distance from yourself and your peer group might help you see that you need to reform your ways or find new friends), it does not appear to be a good thing when we are considering rational demands -- especially if this also includes moral demands. Second, this means that normative alienation provides an *independently* plausible way of characterizing the kind of constraint Korsgaard might be after with Rational Force. According to (vi), at least one necessary condition for Rational Force is that an account of practical reason avoid alienating an agent from the rational demands that apply to her. If normative alienation is intuitively bad, (vi) would appear well supported – independently of whether it suits Korsgaard's own account. A view that satisfies the explanatory demand expressed by (vi) would appear preferable,

¹⁵⁵ See Fitzpatrick (2004) for discussion of normative alienation as an objection to realist views of practical reason. He argues that a Neo-Aristotelian, externalist view of reasons for action is able to avoid the threat of such alienation.

¹⁵⁶ Fitzpatrick (2004), 288.

other things being equal. And, as I will now show, constructivism does appear to do a better job of satisfying this condition than some versions of realism.

For example, compare how our Unrefined Platonist fares with respect to (vi). The Platonist account of reasons is not connected in any *obvious* way with the agent's concerns or the kinds of considerations that figure in her deliberation. On this view, reasons are "out there" in the Platonic heavens. This doesn't mean that such a view necessarily results in normative alienation; rather, it just can't rule this out. To this extent, it does not appear to satisfy the condition in (vi).

Moral constructivism, by contrast, makes reason the outcome of the exercise of our deliberative capacities. Clearly, if reasons are to be explained in terms of deliberation they have a connection to the kinds of considerations that actually figure in an agent's deliberation. What's more, constructivism goes some way toward explaining why someone should cultivate a concern for moral reasons. If moral reasons are the kind of thing that would result from an agent's own sound deliberation, there is a sense in which these reasons are her own.

Now, if all versions of realism accounted for practical reason similarly to an Unrefined Platonism, I do think that (vi) would provide some support for Korsgaard's Assumption. One might think that a view of morality counts as immediately or necessarily connected with an agent's deeply held concerns and interests *only if* it successfully withstands the threat of normative alienation, but that – as in the case of Platonism – no stance-independence view is able to do that. So, if a view is stance-independent, the account of practical reason it provides is not immediately or necessarily connected with an agent's deeply held concerns and interests. The problem

with this justification is that *not* all realist accounts of practical reason work like Platonism.

4.7 Objection: Deliberation Without Construction

In the last section, I suggested that we might understand Rational Force as demanding a satisfactory response to the threat of normative alienation. In this case, Korsgaard's view of practical reason does appear especially successful – and this because of the role it assigns deliberation. But one might ask why constructivism is essential to this kind of response if deliberation is really doing the work. What about the threat of normative alienation requires a specifically constructivist account of practical reason?

Compare, for example, two ways in which an appeal to deliberation might figure in an account of Rational Force understood as (vi):

- A. What it is to be a reason is constrained by facts about us and by the role of deliberation in our lives.
- B. What it is to be a reason is determined by the exercise of our own deliberative capacities.

The second clearly amounts to a constructivist account of practical reason, but the first is compatible with realism. For example, here we might compare the views held by people like Aristotle, Aquinas, Joseph Butler, or T.H. Green. Although there are important differences in the way that each of these philosophers accounts for practical

reason, there are also important similarities. On the one hand, each would appear to endorse something like the realist's commitment to stance-independent moral truth. They think that moral truth is determined in part by facts about our rational nature, and this nature is not ultimately up to us. On the other hand, each would appear to accept something like the view expressed by A. If a commitment to A is sufficient to satisfy (vi), we have no reason to accept Korsgaard's assumption.

So, why isn't a view that incorporates A sufficient as a response to the threat of alienation? According to A, reasons are intimately connected with an agent's capacity for rational deliberation. This is because they are grounded, at least in part, in the very facts that make us what we are and, importantly, this *includes* facts about our deliberative capacities. But, in this case, reasons don't present us with alien demands – quite the contrary! What could present a more immediate connection with our concerns and interests than the very features that make us what we are? But if this is right, there is a kind of stance-independent view of practical reason that can avoid the threat of normative alienation and, consequently, present an immediate or necessary connection with an agent's deeply held concerns and interest. If this is correct, then Korsgaard's Assumption is mistaken and *The Argument from Moral Reasons* unsound. It would appear as if realists as well as constructivists can get (i)-(iii).

In response, Korsgaard would, of course, reject the claim that A alone is sufficient to satisfy (vi). The argument she would give for this claim might go something like the following: Although alienation is surely part of the story, she might say, it is not the *entire* story! In order to account for Rational Force one must *not only* avoid the threat of normative alienation *but also* provide some positive account of why

we identify with our own practical reason. An appeal to A might help us get some of the way, but it is not enough. Rather, what we need is an account of practical reason that is compatible with autonomy. And autonomy, importantly, requires that we self-legislate the very demands that apply to us. This is what *explains* the fact that we are not alienated by them. This is something that B addresses, but not A alone. But since self-legislation is incompatible with a stance-independent account of practical reason, no such account will satisfy autonomy. Consequently, no such view does enough to secure an immediate or necessary connection between an agent's reasons and her deeply held concerns and interests.

In Chapter 5, I will argue at length that this argument from autonomy does *not* end up working. This is because it relies on what I show to be a suspect claim – viz. that autonomy requires self-legislation. But, the success or failure of this argument *notwithstanding*, I think that we can see that we are getting pretty far afield of our three initial considerations and the *Argument from Moral Reasons*.

What I've tried to show in the last part of this discussion is that there is at least a *prima facie* case to be made that Korsgaard's Assumption is mistaken. This is because a realist can secure an intimate connection with deliberation in the way that A describes and, without further argument to the contrary, this looks as it might just be enough to satisfy (vi). If this is right, (vi) alone is too weak for Korsgaard's needs. It *is* a condition that her own view satisfies. What's more, it's also one that rules out some versions of realism – viz., an Unrefined Platonism. But it does *not* rule out all versions of realism. Hence, it does not support premise 6 of our original argument. In this case, we have no good reason to think that constructivism bears a comparative advantage

over realism when it comes to accommodating Common Sense, Rationalism, and Rational Force.

4.8 Conclusion

In section 1, I started out with three very plausible considerations that one might take as constraints on our moral theorizing. I then presented Korsgaard's argument for the claim that moral constructivism does a better job of accommodating these considerations than realism. I showed that premise 6, what I've called *Korsgaard's Assumption*, looked like the most vulnerable point in this argument. The rest of the discussion then involved presenting various conditions for Rational Force that might support this assumption. But these refinements either turned out to be too strong (i.e., Korsgaard's own view failed to satisfy them) or too weak (i.e., in this case, the condition did not obviously rule out all versions of realism). Consequently, it would appear as if the argument against robust moral realism fails. Despite this failure, however, I think that this discussion reveals an important consideration, one – if nothing else – that should influence intramural debates amongst realists. This is that the success of particular species of realism should be evaluated against the threat of normative alienation. As we've seen, this is harder to square with the morality of Common Sense than one might have thought.

CHAPTER 5

AUTONOMY AND THE NATURE OF MORALITY

5.1 Introduction

What implications, if any, does an adequate conception of moral autonomy hold for a metaethical theory? Some philosophers have argued that the centrality of autonomy to moral thought entails a rejection of robust forms of realism – e.g., Rawls (1980) and, in particular, Korsgaard (1996), (2009). This is because, on their view, robustly realist views cannot accommodate the conception of agent-centered control that a plausible account of autonomy requires. This argument, I claim, depends on an account of autonomy that is either question-begging (since the account itself appears to rest on the assumption that there are no choice-independent moral standards) or mistaken (since otherwise it turns out to rest on an important equivocation). I argue, more generally, that the failure of this argument casts significant doubt on any attempt to draw anti-realist conclusions from an adequate conception of autonomy.

Autonomy, it is widely agreed, is essential to the way we conceive ourselves as moral agents. Much of this chapter will be concerned with making this claim more precise, but, very roughly, the understanding of autonomy I have in mind is a *Kantian* one. On this view, autonomy is not *merely* a psychological characteristic that we may possess or lack.¹⁵⁷ Nor is it simply a right we may claim in our interactions with

¹⁵⁷ Here I have in mind the notion of autonomy as a descriptive property had by some or most human beings. For example, I take it that debates about whether a wife who defers to her husband in all decisions counts as autonomous concerns autonomy in this sense. It is a debate about the conditions one

others (though it might imply such a right). Rather, autonomy on this view is an essential characteristic of moral agency. As I shall understand it, to be autonomous is to have and exercise the kind of control over who we are and what we do that is required to be a moral agent.

The question I want to pursue, in particular, is what is supposed to follow from this. One thought is that the notion of autonomy, properly understood, constrains the reasons and obligations that apply to us as moral agents. I'm going to assume with a number of commentators that a minimal requirement of a good moral theory is that it at least be compatible with this widespread idea about autonomy. I'm going to assume, further, that a moral theory that actually supports and explains these claims about autonomy will be attractive as such.¹⁵⁸

If some moral theories have better resources for accounting for autonomy than others, then, to the extent that they do, this will matter to an overall evaluation of their success. Some philosophers have argued that *moral realism* has a particular problem accommodating autonomy and that this gives us reason to reject it. This line of

must satisfy in order to count as exhibiting the psychological condition of self-government. For the origin of this particular example see Hill (1973). Importantly, the notion of autonomy at play in debates between constructivists and realists is not this one. For discussion of these distinctions see Christman (1989), 6ff; see also Feinberg's essay in the same volume (1989), 28ff.

¹⁵⁸ Although the debate between constructivists and realists concerns whether realism is *compatible* with a specific requirement of autonomy, viz., self-legislation, the second point about explanation has also featured in constructivist arguments against realism. In these other arguments, the issue is not so much any particular requirement of autonomy but, rather, the idea that a general characterization of agency and moral personality should play an essential role in a moral theory. John Rawls (1980) concedes that a realist, specifically a rational intuitionist, may admit a role for agency and moral personality in her theory, but only an instrumental one. For example, the intuitionist may assign certain features of moral personality – e.g., impartiality, a disposition to cool-headed reflection, sensitivity to certain features of a situation – a heuristic role. So, the issue is not whether realism is compatible with an account of agency or moral personality. But, Rawls argues that this is not enough. On his view, moral personality is central to our understanding of obligation and moral concern. Hence, a theory, like constructivism, that assigns a characterization of moral personality a central role in the theory will do a better job of accounting for the data.

argument, or something like it, is at least implicit in the work of John Rawls and Christine Korsgaard. We might reconstruct it as follows:

Argument from Autonomy

- (1) Autonomy is essential to the way we conceive of ourselves as moral agents.
- (2) Autonomy requires self-legislation [SL].
- (3) SL is incompatible with robust moral realism.
- (4) So robust moral realism is false.¹⁵⁹

Here we have an argument that begins from the conception of autonomy I have sketched.

Premise (2) states that, so understood, autonomy requires that we self-legislate the terms of morality. In what follows, I will sometimes refer to this as ‘the self-legislation requirement’. It says we stand in an appropriate relation of control to our actions only when we create or construct (in some sense) the standards which determine whether we act well or badly. It is not easy to find a precise formulation of self-legislation in the literature on constructivism. Rawls sometimes describes it as

¹⁵⁹ I am simplifying things a bit here to help focus the discussion in this paper. Constructivists present the *Argument from Autonomy* as part of a package of considerations that allegedly count against realism and in favor of their own view. They then argue on holistic grounds that realism should be rejected. Even if realism can accommodate some of these considerations to some extent, the constructivist argues that her view does a better job on the whole. Hence, it may be more accurate to reword the conclusion to read: *other things being equal, we have good reason to reject realism*. Of course, whether this is the case depends on whether other things *are* equal. For example, a realist might concede that her view loses out to constructivism with respect to autonomy but still insist that other features of realism give it an advantage on the whole. Constructivists might agree that other things are not equal – but contend that this is because constructivism may claim a series of explanatory advantages with respect to other putatively deep features of morality. Adjudicating between these claims will require us to look at each area of claimed comparative strength in turn. This is the task I set out for myself in my dissertation.

the idea that rational agents are “self-originating sources of valid claims.”¹⁶⁰

According to Korsgaard, SL is the view that “...nothing is a law to you except what you make a law for yourself.”¹⁶¹ Neither formulation, however, is as precise as one might hope. Let us understand SL as the claim that *all moral standards are ultimately created by us*. As I will soon explain, this gloss has the advantage of making the argument valid. The challenge I wish to press is whether (2), as such, is true, and the argument sound.

The reason this matters is that, if true, (2) has surprising consequences for the nature of morality. As premise (3) states, SL is incompatible with robust moral realism. Recall that what distinguishes robust realists is their commitment to condition (4). They think that the moral standards that fix our moral principles are not made true by virtue of their ratification from within any actual or hypothetical perspective. In other words, true moral standards are not of our own making. That is, they don't depend on our choices, or beliefs, or desires, etc. But according to premise (2) moral standards must ultimately be of our own making if we are to count as autonomous. So, robust moral realism turns out to be false.

Constructivists, like Korsgaard and Rawls, present this conclusion in the framework of a larger argument that is supposed to support their own considered view. Whereas realism allegedly fails to satisfy the self-legislation requirement, these constructivists contend that their own views succeed. On the one hand, it is easy to see why they would draw this conclusion. Constructivists think that there are many moral principles whose truth is stance-dependent. To the extent that these truths depend on

¹⁶⁰ Rawls (1980), 331ff

¹⁶¹ Korsgaard (2009), 7.5.1

the outcomes of our autonomous deliberation, constructivism may be said to satisfy the self-legislation requirement. On the other hand, the fact that substantive constructivisms, like the ones that Rawls and Korsgaard defend, rely on some small number of stance-independent ethical truths to characterize the moral standpoints which generate these outcomes, one might still ask whether these views are not themselves subject to the objection embodied in the Argument from Autonomy. This is because the stance-independent truths that they appeal to in their characterizations of the moral standpoint are not themselves self-legislated. Although there are moves available to the constructivist in response to this kind of objection, they need not concern us here.¹⁶² For, as I will argue, there is no good reason for us to accept the conclusion to the first stage of the argument.

Of the argument's three premises, (2) is the one that is most vulnerable to criticism. It would also appear to be doing much of the work in the argument. In what follows, I will present and evaluate an argument that Christine Korsgaard has given in defense of premise (2). Besides representing perhaps the only explicit, worked-out defense of the self-legislation requirement in the vast literature on autonomy, Korsgaard's argument appeals to commonly held and quite plausible views to generate a surprising conclusion. This makes her argument both important and interesting.

However, I will show that Korsgaard's argument ultimately rests on a mistake. That is, some of the premises in this argument only make sense if we take them to be referring to autonomy in some other sense than SL. What's more, in order for these

¹⁶² For example, the constructivist might respond that *even if* her view does not fully satisfy the self-legislation requirement, she still does better than the realist. The fact that, according to constructivism, *most* moral truths are the outcomes of autonomous deliberation is still better than a view like realism, according to which very few moral truths work this way.

claims to come out true, SL has to be false. Hence, there is no good reason for us to accept premise (2) and with it the conclusion to the *Argument from Autonomy*.

5.2 Some Preliminaries

As I stated in the beginning, I'm going to assume that premise (1) is true. I'm also going to stipulate a reading of SL that is incompatible with realism. There is the option of leaving the proper interpretation of the term 'self-legislation' open.¹⁶³ For example, this is the line that is sometimes taken in the Kant literature. I'm not doing this; rather, in the interest of simplicity, I will stipulate that we understand SL in the sense that precludes realism. If SL is stipulated in this way, premise (3) is uncontroversially true. This means that the real question to ask is whether premise (2) is true. Does autonomy require SL?

By contrast with other views of autonomy (e.g., autonomy as integrity, authenticity, or self-command), self-legislation, arguably, does not wear its plausibility on its face. Although it clearly bears some connection with the notions of control and self-governance that are at the core of our concept of autonomy, it also represents a sophisticated philosophical claim that would appear to operate at some level of remove from it. In fact, it is very difficult to see what the argument for premise (2) is supposed to be.

¹⁶³ In this case, we allow that on certain readings of 'self-legislation' premise (3) might turn out to be false.

Some philosophers have interpreted Kant as providing such an argument.¹⁶⁴ But, this is extremely controversial. Within Kant scholarship, people have argued that SL cannot be Kant's view of autonomy – on the basis of textual evidence and internal coherence.¹⁶⁵ For the purposes of this paper, I am not interested in what Kant said; rather, I wish to examine independent arguments for SL.

I want to identify an argument for premise (2) that may be extracted from the Kantian apparatus, one that stands on its own. The following methodological consideration will guide my approach: a moral theory that achieves a certain explanatory power with the fewest substantive commitments is preferable, other things being equal. If our acceptance of SL depended on our acceptance of Kant's complicated account of transcendental idealism or transcendental freedom, this would be an unattractive result. This is not because we must assume that Kant is wrong – but rather because Kant's systematic philosophy involves many controversial, substantive commitments (many of which are subject to deep interpretive dispute).

There are considerations that speak both in favor of and against this approach. On the one hand, Karl Ameriks (2003) suggests that interpretations which attribute SL to Kant are driven more by the attractiveness of the substantive view than the evidence. This suggests that it may be possible to find an argument for SL that is independent of the commitments of the Kantian project as a whole. On the other hand,

¹⁶⁴ See, e.g., John Rawls (1980), (2000); Christine Korsgaard (1989), (1996), (2009); Jerome Schneewind (1991), (1998); and Andrews Reath (1994).

¹⁶⁵ See, e.g., Karl Ameriks (2000), (2003); Patrick Kain (2004), (2006); Allen Wood (1999); and Terence Irwin (2004).

Onora O'Neill (1989) appears skeptical that such an extraction is possible.¹⁶⁶ If Ameriks is right, it would seem as if we should bear a good chance of finding independent arguments for premise (2). If O'Neil is right, this task may prove difficult. But, even if this is the case, the result would still be interesting. For it would reveal a weakness of arguments for SL.

5.3 What is Korsgaard's argument for SL?

Christine Korsgaard has presented the most sophisticated, sustained defense of the self-legislation requirement that exists in the literature. Though she presents it in a number of her works, the clearest statements of this argument appear in two of her books, *The Sources of Normativity* (1996) and *Self-Constitution: Agency, Identity, and Integrity* (2009).¹⁶⁷ Although Korsgaard does not name the argument as such, I will refer to it throughout as the *Argument for Self-Legislation*. Here I cite a passage from each book that provides a sketch of the entire argument.

The reflective structure of human consciousness sets us a problem. Reflective distance from our impulses makes it both possible and necessary to decide which ones we will act on: it forces us to act for reasons. At the same time, and relatedly, it forces us to have a conception of our own identity, a conception

¹⁶⁶ Cf. O'Neill (1989): "Those who want to show that autonomy is integral to morality and rationally justified may find that in the end they cannot duck the larger systematic and metaphysical issues" (p. 77).

¹⁶⁷ In her most recent book, Korsgaard is occupied with the issues of personal identity, agency, and action. However, it is clear that she intends the views she presents there to be continuous with, and supplementary to, her earlier work.

which identifies us with the source of those reasons. In this way, it makes us laws to ourselves. When an impulse – say a desire – presents itself to us, we ask whether the maxim of acting on it can be willed as a law by a being with the identity in question. If it can be willed as a law it is a reason, for it has an intrinsically normative structure. If it cannot be willed as a law, we must reject it, and in that case we get obligation.¹⁶⁸

Self-consciousness opens up a space between the incentive and the response, a space of what I call reflective distance. It is within the space of reflective distance that the question whether our incentives give reasons arises. In order to answer that question, we need principles, which determine what we are to count as reasons. Our rational principles then replace our instincts – they will tell us what is an appropriate response to what, what makes what worth doing, what the situation calls for. And so it is in the space of reflective distance, in the internal world created by self-consciousness, that reason is born.¹⁶⁹

Simplifying greatly, the idea is just this. Unlike non-rational animals, which are guided immediately by their instincts (or as Korsgaard puts it at one point, whose instincts constitute their wills),¹⁷⁰ we – as moral agents – are faced with the problem of self-consciousness. Our rational nature allows us to stand back from the flow of impulses that rise up in us and see ourselves as something over and above them. This

¹⁶⁸ Korsgaard (1996), 113

¹⁶⁹ Korsgaard (2009), 6.1.7

¹⁷⁰ Korsgaard (2009): “A non-human animal acts on what I called ‘instinct.’ Her instincts are her principles, and they constitute her will” (10.2.1).

capacity to see ourselves in this way poses a problem. It requires that we *choose* what to do, a need that does not exist prior to reflection.

However, in order for our movements to count as actions, we cannot just choose to follow one of our impulses. If we did that, we wouldn't be acting; rather, the impulse or part would be operating on us. And, in this case, one would be a mere *heap* of impulses, not an agent.¹⁷¹ Instead, we must formulate principles of action that both unify these disparate parts and give us direction. Hence, self-legislation is our response to the task that reflection presents us with. If the problem is that our reflection creates a need for principles or standards to guide us in action, the solution is that we must create these principles or standards for ourselves.

In short: Reflection creates the conditions for moral agency. Moral agency requires that we unify ourselves by giving ourselves standards that make us into someone rather than some thing, standards that do not exist prior to the choices we make in the face of reflective self-consciousness. Hence, in order to act, we must create our own moral standards – i.e., we must self-legislate. Although I think this argument is ultimately flawed, it is not *obviously* flawed.

Part of the appeal of Korsgaard's account of self-legislation comes in the way that it incorporates an independently plausible view of autonomy: *integrity*. Integrity arguably represents a core notion of autonomy. Continually developing or even sometimes reinventing oneself may be good or even required, but someone who reinvented herself from day to day, or moment to moment, or who committed herself

¹⁷¹ As Korsgaard puts it, "to regard some movement of my mind or my body as my action, I must see it as an expression of my self as a whole, rather than as a product of some force that is at work on me or in me. Movements that result from forces working on me or in me constitute things that happen to me" ((2009), 1.4.1).

to radically conflicting and incommensurable ends would fail to be a person, to be a someone. What's more, we think integrity matters morally. We sometimes think that people of deep conviction should act with integrity even if we think they are wrong. We think, for example, that the pacifist should not go to war even if it turns out that the war is just. This view of autonomy may be attractive and the problem of reflective consciousness a real one. But must we think that this requires self-legislation?

5.4 My Strategy for Evaluating the Argument

For the purpose of argument, I will grant that self-consciousness has this destabilizing or fracturing effect on the agent and that this creates a practical problem, one whose solution requires – *in some sense* – that we construct coherent practical identities that we make effective in our actions. However, I will argue that, even if this is all true, it needn't follow that action or agency requires SL. That is, moral agency does not require that we make our own moral standards. This is because there is a basic problem with Korsgaard's *Argument for Self-legislation*.

One initial problem is that, as stated, the argument is invalid. There is an obvious way of fixing this. However, as I will argue, it would require presupposing that moral realism is false. And this is not something that Korsgaard is in a position to do, not if this same claim is the conclusion to the *Argument from Autonomy*. Although Korsgaard suggests an expansion of the original argument for SL that avoids this objection, I will show that this new piece to the argument ultimately rests on an equivocation. In short, if the *Argument for Self-legislation* is to support premise two of

the *Argument from Autonomy* it must conclude that autonomy requires SL. But the crucial step in the former argument turns out to rest on a different notion of autonomy, one that supports our judgments about moral responsibility, not SL.

Again, I should point out that I don't think that Korsgaard's argument is *obviously* flawed. Rather, its problems are obscured by the many details and steps involved. In order to make my objection clear, my approach will be to work backwards. I will start by highlighting a claim that Korsgaard cannot be making. I will then work towards the claim that I think she actually holds. Although there is some distance between these claims, it will be instructive to see how Korsgaard might maintain this distance. Once I have established this much, I can begin to explain where the problem lies – specifically, where she equivocates.

5.5 Reconstructing Korsgaard's Argument

It is important to note a basic distinction: one between (a) choosing what to do and (b) choosing what the moral standards are. These things are clearly different. But might the one entail the other? In other words, does the fact that an agent chooses what to do entail that she also chooses what the moral standards are?

In order to answer this question, let us consider a view on which it is true that an agent chooses what to do but false that she chooses what the moral standards are. Since it is only (b) that causes a problem for the realist, such an alternative would be compatible with moral realism. To illustrate this, let's consider what the story of reflective self-consciousness looks like when we keep this distinction in view. On this

account, when presented with the need to act in reflective self-consciousness, we are forced to find a principle of action that speaks for all the parts of us operating as a whole; it must be a principle of integrity that we think will give our lives meaning and make them worth living. So far, we are working with the same picture that Korsgaard describes. But here's where things diverge.

Unlike Korsgaard's picture, this one allows that there are some important facts about our nature as moral agents – i.e., facts about what kind of practical identities we may coherently form and which may give value to our lives – that exist prior to and independent of our choice. According to this view, the facts might support certain kinds of lives over others. A life of a teacher or a devoted parent might better suit us than the life of a grass-counter or an addict.¹⁷² Autonomous action – that is acting as an agent – requires the appropriate understanding of these facts and the reasons they give us. However, these are facts that we must not only recognize but also *act on*. These facts might determine the standards by which we evaluate whether we act rightly or wrongly, or live good or bad lives. But, nonetheless, we still have to act and live our lives. The moral facts don't live them for us. So, moral agency may require us to choose what to do and who we will be, but this does not require that we create our own moral standards. So, (a) alone does not entail (b).

But perhaps there is some premise that we could add to (a) that, together with it, would entail (b). One obvious candidate is the claim that *there are no moral standards prior to choice*. If we had to choose, but there were no standards available to guide us, our only choice would be to create them, to self-legislate.

¹⁷² However, the facts needn't be fully determinant. For example, they needn't determine whether the teacher's life or the parent's life is better.

But, of course, this is a claim that is not available to Korsgaard. Recall that, in the *Argument from Autonomy*, SL is employed to secure the claim that moral realism is false, or at least incompatible with autonomy. The argument is supposed to show that autonomy is incompatible with prior, or choice-independent, moral standards because autonomy requires SL. But if this is right, the *Argument for Self-legislation* cannot itself rely on the premise that there are no prior moral standards. This would make her overall argumentative strategy circular.

Is there some independent reason to think that moral agency or autonomy is itself incompatible with the prior existence of moral standards? Here is where the details of Korsgaard's more recent work becomes relevant.

Korsgaard's view is not merely that there are no moral standards prior to choice; rather, it is that there are no fixed facts about our nature that may serve to determine these standards. If Korsgaard simply stated that there were no choice-independent moral standards, she would be guilty of the circularity I've just described. But she makes two further claims: moral standards must be grounded in standards of agency, but there are no choice-independent standards of agency to guide us. And, as I'll explain in a moment, this is supposed to make a difference.

I just described a case where self-consciousness forces us to choose what to do and what kind of person to become but this does not force us to accept SL. The reason it does not is that it might still turn out that moral standards are grounded in choice-independent standards of agency. These standards are choice-independent because they are determined by facts about our nature as human beings. Here the facts determine whether a particular practical identity is good or worthwhile, but they don't

live our lives for us. If there were such facts, the realist could claim that the standards they determine provide us with guidance in how to act. The constructivist, however, might simply deny this. She could claim that there are no such facts about our nature that could serve to underwrite choice-independent standards of agency. This would appear to be Korsgaard's response to my proposed alternative.

The view I described earlier states that there are facts about our nature as agents that are substantive enough to guide choice and underwrite morality, but, according to Korsgaard, there are no such facts.

For movement to be my action ... it must result from my entire nature working as an integrated whole... Now this is where things get complicated. You might suppose that this requires an action to be the effect or result or expression of a *prior* unity in the agent, an integrity already achieved... But I will argue that this cannot be how it works. This is where the problem of personal identity comes into the picture. I am going to argue that in the relevant sense there is no *you* prior to your choices and actions, because your identity is in a quite literal way *constituted* by your choices and actions.¹⁷³

¹⁷³ Korsgaard (2009), 1.4.2. See also,

... according to the above argument it is the claim to universality that *gives* me a will, that makes my will distinguishable from the operation of desires and impulses in me. If I change my mind and my will every time I have a new impulse, then I don't really have an active mind or will at all – I am just a kind of location where these impulses are at play... The active will is brought into existence by every moment of reflection, but without the claim to universality, it is no sooner born than dead. And that means that it does not really exist at all. ((1996), 232).

Simply put: she thinks that there are no prior facts about our nature that could serve to guide our choices in reflective self-consciousness. According to Korsgaard, this could only be the case if there were “some *prior* unity in the agent, an integrity already achieved”, but “there is no *you* prior to your choices and actions.”

It is important to note that ‘self-legislation’ is meant quite literally. The picture here is *not* of an agent, a self, standing around thinking about what to do and then coming up with principles to guide herself in action; rather, the picture is one of self-making through lawgiving. Korsgaard claims that we not merely give ourselves laws through acts of willing; rather, we create ourselves – we legislate *the self* into existence.

The function of the normative principles of the will in particular is to bring integrity and therefore unity – and therefore, really, existence – to the acting self.¹⁷⁴ ... It is a law that I give to myself, and its function is to bring unity to myself.¹⁷⁵

In her most recent work, she refers to this process as one of *self-constitution*.

The task of self-constitution involves finding some roles and fulfilling them with integrity and dedication. It also involves integrating those roles into a

¹⁷⁴ Korsgaard (1996), 229

¹⁷⁵ Korsgaard (1996), 230

single identity, into a coherent life...The principles of practical reason bind us because, having to act, we must constitute ourselves as unified agents.¹⁷⁶

Korsgaard claims that we only act well when we stand in the appropriate relation of control to the principles on which we act.¹⁷⁷ But, importantly, what counts as acting well depends on what one is. In other words, moral standards are grounded in standards of agency. But there are no standards of agency prior to choice in the face of reflective self-consciousness. Facts about us independent of our choices cannot determine such standards. Rather, in creating, or constructing, a coherent system of action-guiding principles, we – in the very same act – create or construct a practical identity, and with it the integrity and command over ourselves that is essential to our agency.

Before presenting Korsgaard's reason for thinking that there is no standard-bearing sense of *you* prior to your choices and actions, I should first point out how this general type of response improves on the last one. One might worry that this response runs into the same problem I described in the previous section: it merely ends up

¹⁷⁶ Korsgaard (2009), 1.4.8

¹⁷⁷ One will note that surprising things follow from this conclusion. For example, on this view, the addict does not count as performing *actions*. Korsgaard not only recognizes this consequence, she embraces it. On her view, the addict fails to perform actions because she fails to be an agent; in making herself a slave of one or more of her desires, the addict makes herself into a mere thing, an object – not an agent. Here, one might worry that this means that the addict is not responsible for what she does or the abominable condition she is in. However, this is not Korsgaard's view. She thinks that we are responsible for making ourselves into agents. This is something that we can do well or poorly. Insofar as one fails at making herself into an agent or fails to do this well (and, importantly, external conditions didn't make it impossible for one to do so – e.g., brainwashing), one is subject to blame. The connection between autonomy and responsibility turns out to be central to my objection to the argument. I will discuss it in more detail later in this chapter.

presupposing what Korsgaard must prove. But this worry turns out to be misplaced in this case.

To see what I mean, consider a further distinction: one between (c) choosing what kind of agent to be, and (d) choosing what the standards of agency are. As in the case of the first pair, these two things are also clearly different. And as my example has already shown, (c) alone does not entail (d). The case I have described is exactly one where moral agency may require us to choose what to do and who we will be, but this does not require that we choose what the standards of agency are.

Again, there is a premise that we could add to (c) that, together with it, would entail (d). Here, the obvious candidate is the claim that *there are no substantive standards of agency that exist prior to choice*. If we had to choose what kind of person to be, but there were no prior standards available to guide us, our only choice would be to create them. This appears to be the line that Korsgaard in fact takes, as evidenced by the passage I've cited. But does this commit Korsgaard to the same kind of circularity that was the subject of the previous objection?

In this case, the objection does not appear to hold. The worry in the earlier case was that the *Argument for Self-legislation* tacitly relies on the assumption that there are no choice-independent moral standards, in which case an appeal to the self-legislation requirement in the *Argument from Autonomy* would be circular. However, in this case, the claim assumed is not that there are no choice-independent moral standards; rather, it is that there are no choice-independent standards of agency. It is only in conjunction with the independent claim that the moral standards are grounded

in standards of agency that we get the claim that there are no choice-independent moral standards.

To deny that there are prior facts about our nature that may serve to guide choice is not obviously to deny that there are choice-independent moral standards. This conclusion only follows if we accept the claim that the moral standards are grounded in standards of agency. Although the version of realism I sketched out above follows Korsgaard in grounding standards of morality in standards of agency, a realist needn't do this. So, to deny that there are such facts about agency is not the same as denying choice-independent moral standards. The charge of circularity is averted. But why should we think that there are no choice-independent standards of agency?

5.6 Moral Agency and Responsibility

Korsgaard argues that if there were choice-independent standards of agency we could not rightly be counted responsible for our actions. In other words, moral responsibility requires that we determine the standards of our own agency.

Sometimes you hear philosophers say that the idea of responsibility is incoherent, because we could not be responsible for what we do unless we are responsible for what we are, and we could not be responsible for what we are unless we created ourselves. I think it is true that we could not rightly be held responsible unless we created ourselves, but false that that makes the idea of responsibility incoherent.... But the sense in which we must create ourselves in

order to be responsible is not that we must literally bring ourselves into being. Rather, we are responsible because we have a form of identity that is *constituted* by our chosen actions. We are responsible for our actions not because they are our products but because they are us, because we are what we do.¹⁷⁸

The argument expressed in this passage is compressed. But it appears to go something like this. If there were choice-independent standards of agency, we couldn't freely or autonomously create ourselves. And if we couldn't freely or autonomously create ourselves, we could not rightly be held responsible for our actions. But we can be rightly held responsible for our actions! So, we can autonomously create ourselves. So, there are no choice-independent standards of agency.

After briefly justifying my reconstruction of the argument, I will entertain possible objections to it. Ultimately, however, my evaluation of it will matter less than what the argument shows about Korsgaard's overall conception of autonomy. I will go on to argue in the next section that Korsgaard's appeal to responsibility reveals an

¹⁷⁸ Korsgaard (2009), 6.4.7. These claims are repeated at different points in (2009). For example, see also the following passage from an earlier chapter:

When you deliberately decide what sorts of effects you will bring about in the world, you are also deliberately deciding what sort of a cause you will be. And that means you are deciding who you are. So we are each faced with the task of constructing a peculiar, individual kind of identity – personal or practical identity – that the other animals lack. It is this sort of identity that makes sense of our practice of holding people responsible, and of the kinds of personal relationships that depend on that practice. You will already see that I think those who claim that judgments of responsibility don't really make sense unless people create themselves are absolutely right – only unlike most people who believe this, I don't think it's a *problem*. It is as the possessor of personal or practical identity that you are the author of your actions, and responsible for them. And yet at the same time it is in choosing your actions that you create that identity. What this means is that you constitute yourself *as* author of your actions in the very act of choosing them. (1.4.3)

equivocation in the sense she assigns to *SL*. However, it turns out that this equivocation cannot be reconciled by merely reassigning terms and disciplining how they are used; rather, it reveals a deep confusion in Korsgaard's argument. But before we get this far, let me first say a few words about my presentation of the argument and possible objections that one might level against it.

One will note that I have supplemented the argument in various ways. First of all, the passage I have cited does not explicitly make the connection between the standards of agency and autonomy, or autonomy and responsibility. Rather, what we get is just something like the claim that *if we did not create ourselves, we could not rightly be held responsible for our actions*. It is a common, though not uncontroversial, consideration that responsibility requires that we be free or autonomous in some sense. I take it that this is the kind of consideration behind the above conditional. Specifically, the claim is that responsibility requires autonomy in the sense of self-creation, or what Korsgaard calls *self-constitution* in other places. In other words, if we did not self-constitute, we couldn't rightly be held responsible for our actions.

But if one accepts this claim, one might also think that choice-independent standards of agency would somehow infringe on or preclude this kind of freedom or autonomy. Clearly, if there are choice-independent standards of agency, then we do not create these standards ourselves by choosing them. This is just what it means for a standard to be *choice-independent*. This leaves unresolved, however, the question of whether self-constitution requires that we make or create these standards. Korsgaard makes it clear here and elsewhere that she takes self-constitution to require this. It is

not merely her view that we choose practical identities and what kind of roles we will organize our lives around; rather, it is that in doing so an agent determines the standards of her own agency – i.e., standards which determine whether she succeeds or fails *as an agent*.¹⁷⁹ In other words, on her view, if there were choice-independent standards already there, we could not be said to self-constitute in the sense she takes to matter for moral responsibility. Although Korsgaard does not spell these claims out explicitly, they are consistent with her presentation and would appear to enjoy support in this context.

Furthermore, nowhere in the passage above does Korsgaard explicitly state that we can rightly be held responsible for our actions.¹⁸⁰ I take it that this must be a suppressed premise. Otherwise, it is difficult to see how the argument goes through. Again, this is not an uncontroversial assumption, but without it we could only conclude that *if* there were choice-independent standards for our agency, we could not rightly be held responsible for our actions. But if this is all that the argument shows, and the realist view I described above were correct, the conclusion to draw would be that we are *not* rightly held responsible for our actions. So much the worse for us! But it is clear that Korsgaard wishes to block this conclusion.

¹⁷⁹ She presents an interpretation of this claim against the background of what she takes to be Aristotle's view of the human form: "But the form of the human is precisely the form of the animal that must create its own form" ((2009), 6.4.3). It is clear from what she goes on to say in this context that a *form* constitutes the standards of agency for the being in question.

¹⁸⁰ One could argue that it is entailed by the claim that "...we are responsible because we have a form of identity that is *constituted* by our chosen actions." But it is not clear how we are to read this. Earlier in the passage Korsgaard talks about the conditions that must obtain in order for the *idea of responsibility* to be coherent. In light of this, it may be more accurate to rephrase the above claim as *the idea of responsibility (i.e., the claim that we are responsible) is coherent only if we have a form of identity that is constituted by our actions*. But from this conditional formulation it would be illegitimate to infer that we are in fact responsible.

She introduces the topic of responsibility in order to support the claim that there are no such choice-independent standards of agency. In other words, on her view, the realist picture I have described is mistaken. And in order to secure this claim, she must assume that we can in fact be rightly held responsible for our actions. Some might wish to deny this. But I am happy to grant it. If we can't rightly be held responsible, this would mean attributing massive error to our everyday moral thinking. Perhaps this is a conclusion that argument forces upon us. But the burden is arguably with the person who wishes to attribute massive error to our moral thinking, not with the person who assumes that we can rightly hold people responsible for their actions. In any case, this assumption is probably the least controversial premise in the argument. There is much else in this argument that is also controversial. For example, one might also object to either of the argument's other two premises. In fact, both would appear suspect.

On the one hand, one could deny that if there were choice-independent standards of agency, we couldn't freely or autonomously create ourselves. Of course, Korsgaard stipulates that self-constitution (i.e. self-creation) is incompatible with choice-independent standards of agency. Again, *self-constitution* is a technical term for her. But we might ask whether there is some other non-technical sense in which we might be said to "create ourselves" that does not stand at odds with choice-independent standards of agency. Consider the case I presented earlier. There the point was that *even if* prior standards of agency exist, we still have to live our lives and decide what kind of people we will be. The standards cannot live our lives for us. One might think that deciding what kind of person to become counts as a kind of self-

creation, even if it is not *self-creation* in the technical sense that Korsgaard has in mind, and that this suffices for autonomy. In other words, the mere presence of prior standards would not appear to preclude autonomy – that is, unless we are assuming that autonomy requires SL. And Korsgaard cannot be assuming the self-legislation requirement at this stage in the argument.

Recall that this claim about the standards of agency is supposed to figure in an argument *for* SL. That is, the reason provided for thinking that we must self-legislate is allegedly that there are no prior standards of agency to guide us. This means that we cannot rely on the self-legislation requirement to argue that there are no prior standards. Barring this assumption, there is no reason to think that the existence of such prior standards precludes autonomous self-creation in some other important sense.

On the other hand, one could also deny the claim that if we couldn't freely or autonomously create ourselves, we could not rightly be held responsible for our actions. Korsgaard remarks that some philosophers maintain this view. But she does not reveal their identities or their reasons for thinking this to be the case.¹⁸¹ Again much would appear to depend on how we understand autonomous self-creation.

The sense of autonomy or self-creation in the antecedent cannot be SL for the same reason that it cannot be in the previous premise. The reason Korsgaard is concerned to defend the claim that there are no standards of agency prior to choice is that she takes it that this will somehow help her defense of SL. So, again in this case,

¹⁸¹ Perhaps she has in mind the kind of argument presented by Galen Strawson in “The Impossibility of Moral Responsibility” (1994).

she cannot rely on the view that *autonomy requires SL* in a lemma that is supposed to ultimately support that very conclusion.

One could object that responsibility only requires that we have some freedom in choice about what we do, not who we are. The claim that responsibility requires freedom in choice about what we do is very plausible. But the claim that freedom in choice about what we do requires freedom in choice about what we are is much less plausible. Further to the point, it is not clear that responsibility for what we do requires responsibility for what we are. There might be certain extreme cases where this seems right.¹⁸² But surely I can be responsible for what I do even if I am not responsible for who I am. I am the son of Arrigo and Sandra. This is not an identity that I chose. Additionally, let us suppose that if I had had the choice, I would not have taken this identity; rather, I would have chosen to be born to some other set of parents.¹⁸³ This counterfactual choice notwithstanding, I think that I would still be responsible for whether I make good on my filial obligations to *them*, my actual parents. Along the same lines, we think that one can be a rotten kid to rotten parents. By this I mean that you can be rotten for failing to live up to your filial obligation even if your parents are

¹⁸² For example, if someone is born into a life of extreme poverty and crime this might attenuate our willingness to hold them *morally* responsible for certain actions.

¹⁸³ Mom and Dad, if you are reading this, don't worry! I love you guys. This is just a thought experiment.

rotten, sorry excuses for parents.¹⁸⁴ Families are tough, but you are still responsible for your obligations to them – even if you don't get a choice about this.¹⁸⁵

As this last point and the previous one illustrate, Korsgaard's argument for the claim that there are no choice-independent standards of agency appears vulnerable to criticism in different ways. Of course, for all I have said, it might still turn out that there are good responses available to my objections. But even if there are, this will not save Korsgaard's overall argumentative strategy. This is because Korsgaard's arguments ultimately rely on an equivocation.

We are now in a position to see the entire argument for SL. Let us quickly review. According to Korsgaard, we are autonomous only if we act on self-legislated principles that unify us as something over and above our impulses. In other words, we are autonomous insofar as we live up to our own chosen standards. But the reason for thinking that we must self-legislate these principles, or standards, is that moral standards are grounded in standards of agency, and there are no standards of agency prior to choice. So we have to make our own. Furthermore, the reason that there are no choice-independent standards of agency is that if there were we would lack the kind of

¹⁸⁴ I admit that there may be circumstances in which these obligations are cancelled out by other features of one's situation. For example, if a child is the victim of parental abuse this may cancel her obligation to care for her parent in certain ways – in fact, perhaps this would cancel any kind of filial responsibility to them. But I reject the idea that one's merely failing to identify with one's actual parents, or counterfactually deciding not to be born to them, would count as such a feature.

¹⁸⁵ Korsgaard rejects my claim about filial obligations. She thinks that "...you can walk out even on factually grounded identity like being a certain person's child or a certain nation's citizen, dismissing the reasons and obligations that it gives rise to, because you just don't identify yourself with that role. Then it's not a form of *practical* identity anymore: not a description under which you value yourself" ((2009), 1.4.7). I think that she is mistaken, but I do not have the space to fully argue for this point. My point here is simply to provide some initial reasons for thinking that her argument for the claim that there are no choice-independent standards fails. However, as it turns out, I am less concerned to show the failure of this argument than to highlight what it shows about Korsgaard's general conception of autonomy. Insofar as she appeals to claims about moral responsibility to defend self-constitution, Korsgaard cannot be talking about autonomy as SL.

autonomy required for responsibility. We are responsible only if we are autonomous in the sense that we choose our own standards of agency. But if responsibility requires that we freely choose our own standards of agency, these same standards cannot ground morality as Korsgaard claims.

5.7 Two Incompatible Senses of Autonomy

Now that we have the entire argument before us, we can see how Korsgaard equivocates. Consider these two claims:

- (i) You realize autonomy only if you act according to your self-legislated standards – i.e., insofar as you act well.
- (ii) You realize autonomy in acting badly just as much as acting well – that’s why we rightly hold you responsible for your actions.

Claim (i) follows from the claim that autonomy requires SL. Claim (ii), however, is a natural way of expressing our reason for thinking that responsibility requires autonomy. We want to be able to legitimately impute blame to people for wrong actions. But in order to do this, so the thought goes, they must be responsible. And, as Korsgaard argues, responsibility requires autonomy. But there is a problem. If ‘autonomy’ is to be read univocally, claims (i) and (ii) are inconsistent. It can’t be the

case that you realize autonomy in acting badly AND only if you act well.¹⁸⁶ Either one of these claims is false, or ‘autonomy’ does not have the same meaning in both of them. But either way Korsgaard’s argument fails.

If one of them is false, then Korsgaard’s argument turns out to be unsound. If (i) is false, autonomy does not require SL. In this case, we ought to reject the second premise of the *Argument from Autonomy*, along with its conclusion that realism is false.

If (ii) is false, this is presumably because one is not autonomous in acting badly; rather, as (i) states, it’s only in those cases in which we act well that we realize our autonomy. But, in this case, we have no reason to accept the claim that there are no prior standards of agency to guide us in choice.

Korsgaard’s reason for thinking that we must create our own standards is that if we didn’t we couldn’t rightly be counted as responsible for our actions. But if you aren’t autonomous in those cases where you act badly, then you aren’t responsible for them. This just follows from Korsgaard’s own view that responsibility requires autonomy. But then there is no problem for prior standards. Prior standards might determine that your actions are wrong. But since you are not responsible for them anyway, there is no reason to think that there couldn’t be such choice-independent standards of agency. In other words, this allows room for my realist alternative according to which we still must decide what to do and what kind of person to be *even if* there are standards of agency prior to choice. This is a counterexample to the claim

¹⁸⁶ This is the very same problem that Henry Sidgwick finds in the work of Kant. See “The Kantian Conception of Free Will”, originally printed in *Mind*, 1888, Vol. XIII., No. 51; partially reprinted as an appendix in Sidgwick (1966).

that choosing what to do requires creating your own standards of agency. Now that there is no reason not to accept this realist picture, it would appear as if a crucial premise in Korsgaard's *Argument for Self-legislation* is false. In this case, her argument for SL is unsound. Again, this also means that there is no good reason to accept the conclusion presented in the *Argument from Autonomy*.

If 'autonomy' had a different meaning in (i) and (ii), then both of these claims could be true. But in this case Korsgaard's argument would turn out to be invalid. In order to get the conclusion that realism is false in the *Argument from Autonomy*, one must suppose that autonomy requires SL in the sense that we create our own standards of morality. But the kind of autonomy we get from her *Argument for Self-legislation* is merely autonomy in a sense that is required for responsibility and, importantly, this is a kind of autonomy that is compatible with choice-independent standards of agency. Despite its namesake, this argument does not secure SL. This equivocation means that Korsgaard cannot legitimately rely on the results of the latter argument to support her conclusion in the former.

This is a deep problem for Korsgaard's argument, one that cannot simply be avoided by reassigning terms and regimenting their use. If this is not already clear, it will become so when we consider Korsgaard's own response to the objection and why it fails.

5.8 Why Korsgaard's Response Fails

Korsgaard is aware that there is something like this problem for her overall view. She takes it be the same problem that Kant and Plato face given that, on her reading of their works, both of them accept that we are autonomous only insofar as we realize our “true” selves. But her response ultimately fails to address the objection that I have presented. This is because it only explains how we can be responsible for defective actions, not how we can be responsible for defective characters or practical identities. There is no way of extending her account of defective action to defective character without also rejecting SL in the sense required for the *Argument from Autonomy*.¹⁸⁷

On Korsgaard's view, an unjust or wrong action is one that fails to make one into an agent. This follows because she thinks that moral standards are grounded in standards of agency. Korsgaard recognizes a problem here: “But if this is so, then it is hard to see how we can be held responsible for bad or heteronomous action, or indeed why we should regard it as action at all.”¹⁸⁸ She then devotes an entire chapter of (2009) to addressing it.¹⁸⁹

Part of her response is that actions come in degrees, depending on how well they unify an agent. But responsibility is an all or nothing affair. Insofar as you have some practical identity or other, you can be held responsible for failing to live up to it.

¹⁸⁷ Importantly, my version of the objection would only count as a problem for Plato or Kant if these philosophers also accepted SL. Plato, of course, would reject SL. On his view, whether or not we act well in harmonizing the parts of our souls depends on whether, and the extent to which, one's principles share in the *forms* – and whatever disputes there might be about what this all involves, it is clear that it does *not* involve something like SL. However, it is matter of deep scholarly dispute whether Kant accepts SL. See notes 7 and 8 for some of the relevant literature on this point.

¹⁸⁸ Korsgaard (2009), 8.1.2

¹⁸⁹ See Chapter 8, “Defective Action,” in (2009).

And this makes sense. Consider, for example, the case of G.A. Cohen's idealized

Mafioso:

This Mafioso does not believe in doing unto others as you would have them do unto you: in relieving suffering just because it is suffering, in keeping promises, because they are promises, in telling the truth because it is the truth, and so on. Instead, he lives by a code of strength and honour that matters as much to him as some of the principles as I said he disbelieves in matter to most of us.¹⁹⁰

Let us suppose, as Cohen does, that this Mafioso has the practical identity he does because he has self-constituted himself in a particular way. That is, he meets all of the formal conditions that Korsgaard describes for making oneself into an agent.¹⁹¹ But this doesn't mean that he only has this practical identity in those moments where he lives up to his standards. Presumably, we can only speak of someone having a particular practical identity against the background of many actions. Self-created standards are not born in a single act, nor can they be extinguished by a single act. They enjoy stability and provide a backdrop against which we may judge someone to act well or badly. That is, because of this backdrop, we can judge someone to act

¹⁹⁰ Cohen in Korsgaard (1996), 183

¹⁹¹ It should be noted that Cohen has a somewhat different aim in appealing to this case. He takes it to be a direct counterexample to Korsgaard's view that constructivism can guarantee something like our common-sense morality (i.e., a set of moral standards that include the usual duties of aid and forbearance). It is supposed to be a case where all of the formal conditions of reflective endorsement are satisfied but the resulting moral code is not one that lines up with the morality of common sense. By contrast, I am using this example to draw out the ways in which Korsgaard can explain defective action, but not defective character. Both Cohen's point and mine are related but they are employed to slightly different ends.

badly and – at the same time – hold them responsible for this failing. It is because our Mafioso has a stable and enduring practical identity that we can judge him to act badly when he fails to live up to his own standards – e.g. when he backs down from an act of vengeful violence because he lacks resolve. On this view, the more an agent's actions conform to her own self-created practical identity, the more she can be said to realize autonomy. However, in order to be held responsible for one's actions, all that is required is that one have self-legislated at all. So, Korsgaard's account would appear to allow both that defective action is possible and that we may be rightly held responsible for such actions.

The problem with this response, however, is that it only explains how we can be responsible for defective actions, not defective characters or practical identities. And, importantly, Korsgaard's argument that there are no choice-independent standards of agency relies on the claim that we can only be responsible for what we do if we are responsible for who we are. If you are responsible for who you are, we should be able to impute blame to you for the kind of character you have created for yourself. And this cannot be the case if SL is true in the sense needed to rule out realism.

To say that the Mafioso is responsible for choosing the particular identity he has and not another is to say that we may legitimately praise or blame him for this choice. For this to be the case, however, there would need to be some prior standards of agency against which we could judge him to be blameworthy or not – i.e., some standards that he might fail to live up to. But if SL is true, there cannot be. So, either

the Mafioso is not responsible for choosing who he is, or SL is false. Either way, Korsgaard has failed to adequately respond to the objection I have presented.

If he is not responsible, this would again undermine the case *against* choice-independent standards of agency. Korsgaard's argument for the claim that there are no such standards assumed that we can rightly be held responsible for who we are. If this is false, then the conclusion that there are no choice-independent standards of agency does not follow. In this case, we have no reason to think that we must create our own standards. Instead, we may accept that the realist picture I have provided is correct. We might be forced to choose who we will be, but this does not require that we choose the very standards which determine whether we succeed or fail *as an agent*. On the other hand, if SL is false, the *Argument from Autonomy* is again unsound.

5.9 Conclusion

I started out this discussion with an initial claim: Autonomy is essential to the way we conceive of ourselves as moral agents. My guiding question throughout has been whether there is anything interesting about the nature of morality that is supposed to follow from this.

Although Korsgaard maintains that considerations from autonomy provide reason for thinking that moral realism is mistaken, I have argued that the *Argument for Self-Legislation* does not secure the necessary premise. Korsgaard must provide some good reason to think that autonomy requires SL. But it turns out that her argument for SL relies on a different sense of autonomy, one that is required for legitimate

attributions of responsibility. And, as I have shown, this kind of autonomy stands in deep tension with SL.

Although part of her argument appeals to a plausible view of autonomy in integrity, I have argued that there is a way of accommodating this without SL. Even if we accept that the problem that reflective self-consciousness presents us with is a real one, one that requires that we construct coherent practical identities that we make effective in our actions, this much may be accommodated without its being the case that we create our own moral standards. Autonomy might require that we choose what to do, but this alone does not entail that we make our own moral standards. It does not require SL. Hence, the argument provides no reason to think that robust moral realism is false; there is no reason to reject condition (4). This also means that there is no comparative advantage to be had by a view that grounds most moral truths in the outcomes of autonomous deliberation. The Argument from Autonomy provides no reason for thinking that a substantive constructivism, like the kind that Rawls or Korsgaard defend, is preferable to a robust moral realism.

We are now at the end of our journey. Although I hope it is clear how we might understand the various ways in which constructivism poses a threat to realism, I also hope to have shown that none of these threats proves fatal.

Many have dismissed constructivism on the grounds that it does not appear to present a new or interesting view. This is because they have assumed, quite reasonably, that constructivism is supposed to offer a free-standing alternative to existing metaethical views. But this assumption turns out to be mistaken. I argue that constructivism is better characterized, more restrictedly, as challenge to either the

nature or scope of the realist's conception of ethical objectivity. It is only once we have accepted this characterization that we can see how powerful the constructivist's challenge can be.

Contra current consensus, I hope to have shown that the most powerful version of this challenge is *not* represented by formal presentations of constructivism; rather, it is to be found in the traditional presentations and arguments one encounters in the works of John Rawls, Christine Korsgaard, and T.M. Scanlon. Substantive constructivism may not possess any significant metaphysical or epistemological advantages over realism, but there are good initial reasons to worry that the ethical considerations they marshal in *The Argument from Moral Reasons* and *The Argument from Autonomy* do support constructivist claims to a holistic explanatory advantage over realism.

Although I argue that each of these arguments ultimately proves unsound, they turn out to reveal interesting things about what the right version of realism might look like. For example, the importance of autonomy and the connection between morality and reasons for action suggests that a realist should assign *practical reason* an important role in the human good – though this should be construed in realist terms, not constructivist ones. It also prompts important yet difficult questions about the extent to which a realist may incorporate or accept constructivist elements in her theory.

Much of the preceding discussion has rested on the assumption that condition (4) captures what many defenders of realism are after. Although I have provided some reason for thinking that this condition is what distinguishes realists from substantive

constructivists, there is certainly more work to be done. My aim throughout has been to provide a sketch of a framework that might serve to advance our understanding of the debate. I hope that the arguments in this work mark a step, however small, in this direction.

WORKS CITED

- Ameriks, Karl. 2000. *Kant and the Fate of Autonomy: Problems in the Appropriation of the Critical Philosophy*. Cambridge: Cambridge University Press.
- Ameriks, Karl. 2003. "On Two Non-Realist Interpretations of Kant's Ethics." In *Interpreting Kant's Critiques*. Oxford: Clarendon Press, pp. 263-282.
- Anscombe, Elizabeth. 1957. *Intention*. Oxford: Basil Blackwell.
- Blackburn, Simon. 1984. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Clarendon Press.
- Blackburn, Simon. 1993a. *Essays in Quasi-Realism*. New York: Oxford University Press.
- Blackburn, Simon. 1993b. "Circles, Finks, Smells, and Biconditionals." *Philosophical Perspectives* 7, pp. 259-279.
- Blackburn, Simon. 1998. *Ruling Passions*. Oxford: Oxford University Press.
- Brink, David. 1989. *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press.
- Brink, David. 1992. "A Puzzle About the Rational Authority of Morality." *Philosophical Perspectives* vol. 6, Ethics, pp. 1-26.
- Brink, David. 1993. "Rational Egoism, Self, and Others." In *Identity, Character, and Morality*. Eds. Owen J. Flanagan and Amélie Oksenberg Rorty. Cambridge, MA: MIT Press, pp. 339-378.
- Brink, David. 2001. "Realism, Naturalism, and Moral Semantics." *Social Philosophy & Policy* 18, pp. 154-76.
- Casati, Roberto and Christine Tappolet (eds.). 1998. *European Review of Philosophy* 3: *Response-Dependence*.
- Christman, John. 1989. *The Inner Citadel: Essays on Individual Autonomy*. New York: Oxford University Press.
- Cuneo, Terence. 2001. "Are moral qualities response-dependent?" *Nous* 35, pp. 569-

591.

- Cuneo, Terence. 2007. "Recent Faces of Moral Realism." *Philosophy Compass* 2/6: pp. 850–879.
- Darwall, Stephen, Allen Gibbard, and Peter Railton. 1992. "Toward a Fin de siecle Ethics: Some Trends." *The Philosophical Review* Vol. 101, No. 1, Philosophy in Review: Essays on Contemporary Philosophy (Jan.), pp. 115-189: 137-144.
- Dworkin, Ronald. 1996. "Objectivity and Truth: You'd Better Believe It." *Philosophy & Public Affairs* 25, pp. 87-139.
- Enoch, David. 2007. "An Outline of an Argument for Robust Metanormative Realism." In *Oxford Studies in Metaethics*, vol. 2, R. Shafer-Landau (ed.), Oxford: Oxford University Press, pp. 21-50.
- Enoch, David. 2009. "Can There be a Global, Interesting, Coherent Constructivism about Practical Reason?" *Philosophical Explorations* vol. 12, no. 3, pp. 319-339.
- Enoch, David. 2010. "How Objectivity Matters." In *Oxford Studies in Metaethics* vol 5, R. Shafer-Landau (ed.), Oxford: Oxford University Press, pp. 111-152.
- Feinberg, Joel. 1989. "Autonomy." In Christman, pp. 27-53.
- Finlay, Stephen and Mark Schroeder. 2008. "Reasons for Action: Internal versus External", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = < <http://plato.stanford.edu/entries/reasons-internal-external/> >.
- Fitzpatrick, William. 2004. "Reasons, Value, and Particular Agents: Normative Relevance Without Motivational Internalism." *Mind* vol. 113, issue 450, pp. 285-318.
- Fitzpatrick, William. 2005. "The Practical Turn in Ethical Theory: Korgaard's Constructivism, Realism and the Nature of Normativity." *Ethics* 115, pp. 651-691.
- Fitzpatrick, William. 2008. "Robust Ethical Realism, Non-Naturalism and Normativity." In *Oxford Studies in Metaethics*, vol. 3, R. Shafer-Landau (ed.), Oxford: Oxford University Press, pp. 159-205.
- Foot, Philippa. 1972. "Morality as a System of Hypothetical Imperatives." Reprinted

- in *Virtues and Vices*. Berkeley: University of California Press, 1978, pp. 157-173.
- Foot, Philippa. 1983. "Moral Realism and Moral Dilemma." *Journal of Philosophy* 80, pp. 379-398.
- Gibbard, Allan. 1999. "Morality as Consistency in Living: Korsgaard's Kantian Lectures." *Ethics* 110, pp. 140-164.
- Gibbard, Allan. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
- Harman, Gilbert. 1975. "Moral Relativism Defended." *The Philosophical Review* vol. 84, pp. 3-32.
- Harman, Gilbert. 2001. "Is There a Single True Morality?" In Moser, P. and Carson, T. (eds.), *Moral Relativism: a Reader*. Oxford: Oxford University Press, pp. 165-184.
- Hill, Thomas. 1973. "Servility and Self-Respect." *The Monist* vol. 57, no. 1, pp. 43-61.
- Howard-Snyder, Frances and Alastair Norcross. 1993. "A Consequentialist Case for Rejecting the Right." *Journal of Philosophical Research* vol. 18, issue 0, pp. 110-126.
- Hume, David. 2000. *A Treatise of Human Nature*. Eds. David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.
- Hussain, Nadeem, and Nishi Shah. 2006. "Misunderstanding Metaethics: Korsgaard's Rejection of Realism." In *Oxford Studies in Metaethics* vol 1, R. Shafer-Landau (ed.), Oxford: Oxford University Press, pp. 265-294.
- Irwin, Terence. 2004. "Kantian Autonomy." In *Action and Agency*, ed. H. Steward and J. Hyman. Cambridge: Cambridge University Press.
- James, Aaron. 2007. "Constructivism about Reasons." *Philosophy and Phenomenological Research* 74 no. 2
- Johnston, Mark. 1989. "Dispositional theories of value." *Proceedings of the Aristotelian Society* suppl. vol. 62, pp. 139-174.
- Johnston, Mark. 1993. "Objectivity refigured: Pragmatism without verificationism." In

- J. Haldane and C. Wright (eds.) *Reality, Representation, and Projection*. Oxford: Oxford University Press, pp 85-130.
- Kain, Patrick. 2004. "Self-Legislation in Kant's Moral Philosophy." *Archiv für Geschichte der Philosophie* 86, pp. 257–306.
- Kain, Patrick. 2006. "Realism and Anti-Realism in Kant's Second Critique." *Philosophy Compass* 1/5, pp. 449–465.
- Kant, Immanuel. 1998. *Groundwork of the Metaphysics of Morals*. Trans. by Mary J. Gregor. New York: Cambridge University Press.
- Korsgaard, Christine. 1989. "Morality as Freedom." In *Kant's Practical Philosophy Reconsidered*, ed. Yirmiyahu Yovel. Dordrecht: Kluwer Academic Publishers, pp. 23-48.
- Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Korsgaard, Christine. 1997. "The Normativity of Instrumental Reason." In *Ethics and Practical Reason*, ed. Garrett Cullity and Berys Gaut. Oxford: Clarendon Press, pp. 215-254.
- Korsgaard, Christine. 2003. "Realism and Constructivism in Twentieth-Century Moral Philosophy." *Philosophy in America at the Turn of the Century*, APA Centennial
- Korsgaard, Christine. 2009. *Self-constitution: Agency, Identity, and Integrity*. New York: Oxford University Press.
- Lenman, James. 2010. "Humean Constructivism in Moral Theory." In *Oxford Studies in Metaethics* vol 5, R. Shafer-Landau (ed.), Oxford: Oxford University Press, pp. 175-193.
- Lewis, David. 1989. "Dispositional Theories of Value", *Proceedings of the Aristotelian Society* Suppl. Vol. 63, pp. 113-37.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin Books.
- Milo, Ronald. 1995. "Contractarian Constructivism." *The Journal of Philosophy* 92, pp. 181-204.
- Moore, G.E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.

- Nagel, Thomas. 1986. *The View from Nowhere*. New York: Oxford University Press.
- Norcross, Alastair. 2006 "The Scalar Approach to Utilitarianism." In *The Blackwell Guide to Mill's 'Utilitarianism'*. Ed. Henry West. Blackwell Publishing: Malden, MA, pp. 212-232.
- O'Neill, Onora. 1989. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.
- O'Neill, Onora. 2003. "Constructivism vs. Contractualism," *Ratio (new series)* XVI 4 December, pp. 319-331.
- Parfit, Derek. 1997. "Reasons and Motivation." *Supplementary volume - Aristotelian Society* vol. 71, pp. 99-130.
- Parfit, Derek. 2006. "Normativity." In Russ Shafer-Landau (ed.) *Oxford studies in metaethics* vol. 1, Oxford: Clarendon Press, pp. 325-380.
- Parfit, Derek. Forthcoming. *On What Matters*. Oxford: Oxford University Press.
- Pettit, Philip. 2000. "Review: Two Construals of Scanlon's Contractualism." *The Journal of Philosophy*, vol. 97 no. 3, pp. 148-164.
- Plato. 1997. *Euthyphro*. In *Plato: Complete Works*, John Cooper and D.S. Hutchinson (eds.), G.M.A. Grube (trans.), pp. 1-16
- Pettit, Phillip. 1991. "Realism and response-dependence." *Mind* 100, pp. 587-626.
- Prior, A.N. 1949. *Logic and the Basis of Ethics*. Oxford: Clarendon Press.
- Railton, Peter. 1984. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* Vol. 13, No. 2, pp. 134-171.
- Railton, Peter. 1986. "Moral Realism." *Philosophical Review* Vol. 95, pp. 163-207.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Rawls, John. 1980. "Kantian Constructivism in Moral Theory." *The Journal of Philosophy* 77, pp. 515-572.
- Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.
- Rawls, John. 1999. *Collected Papers*. Samuel Freeman (ed.), Cambridge, MA: Harvard University Press.
- Rawls, John. 2000. *Lectures on the History of Moral Philosophy*. Cambridge: Harvard University Press.

- Reath, Andrews. 1994. "Legislating the Moral Law," *Nous* 28, pp. 435–64.
- Ridge, Michael. Forthcoming. "Kantian Constructivism: Something Old, Something New." In a volume on constructivism for Oxford University Press, ed. James Lenman.
- Sayre-McCord, Geoffrey. 1988. "Introduction: The Many Moral Realisms." In *Essays on Moral Realism*, Ithaca: Cornell University Press. Pp. 1-23.
- Scanlon, T.M. 1982. "Contractualism and Utilitarianism", in A. Sen and B. Williams (eds.), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press), pp. 103-28.
- Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scanlon, T.M. 2003. "Metaphysics and Morals." *Proceedings and Addresses of the American Philosophical Association* vol. 77 no. 2, pp. 7-22.
- Schneewind, Jerome. 1991. "Natural Law, Skepticism, and Methods of Ethics." *Journal of the History of Ideas* 52, pp. 289–308.
- Schneewind, Jerome. 1998. *The Invention of Autonomy: A History of Modern Moral Philosophy*. Cambridge: Cambridge University Press.
- Schroeder, Mark. 2008. *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.
- Shafer-Landau, Russ. 2003. *Moral Realism: A Defence*. Oxford: Clarendon Press.
- Sidgwick, Henry. 1966. "The Kantian Conception of Free Will." Partially reprinted as appendix to *The Methods of Ethics*. New York: Dover Publications, Inc., pp. 511-516. Originally printed in *Mind*, 1888, Vol. XIII., No. 51.
- Smith, Michael. 1989. "Dispositional theories of value." *Proceedings of the Aristotelian Society* suppl. vol. 62, pp. 89-111.
- Smith, Michael. 1993. "Objectivity and moral realism: On the significance of the phenomenology of moral experience." In J. Haldane and C. Wright (eds.) *Reality, Representation, and Projection*. Oxford: Oxford University Press, pp. 235-55.
- Strawson, Galen. 1994. "The Impossibility of Moral Responsibility." *Philosophical*

- Studies* 75, pp. 5-24.
- Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127, pp. 109-166.
- Street, Sharon. 2008. "Constructivism About Reasons." *Oxford Studies in Metaethics* vol. 3, ed. Russ Shafer-Landau. Oxford: Clarendon Press.
- Street, Sharon. 2009. "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues* (a supplement to *Nous*) vol. 19, ed. Ernest Sosa, pp. 273-298.
- Street, Sharon. 2010. "What is Constructivism in Ethics and Metaethics?" *Philosophy Compass* 5/5, pp. 363-384.
- Sturgeon, Nicholas. 2003. "Moore on Ethical Naturalism." *Ethics* 113, pp. 528-56.
- Timmons, Mark. 2003. "The Limits of Moral Constructivism." *Ratio (new series)* XVI 4 December, pp. 391-423.
- Wallace, R. Jay. 2008. "Practical Reason", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = < <http://plato.stanford.edu/entries/practical-reason/>>.
- Wilson, Jessica. 2010. "What is Hume's Dictum, and Why Believe It?*" *Philosophy and Phenomenological Research* vol. LXXX, no. 3, pp. 595-637.
- Wood, Allen. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.
- Wright, Crispin. 1988. "Moral Values, Projections, and Secondary Qualities." *Proceedings of the Aristotelian Society* suppl. vol. 62: 1-26.
- Zangwill, Nick. 2003. "Against moral response-dependence." *Erkenntnis* 59, pp. 285-290.