



# Objectivity and Intersubjectivity in Moral Philosophy

## Link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:40046415>

## Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

## Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#)

**OBJECTIVITY AND INTERSUBJECTIVITY IN MORAL PHILOSOPHY**

A DISSERTATION PRESENTED

BY

PAUL JULIAN

TO

THE DEPARTMENT OF PHILOSOPHY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE SUBJECT OF

PHILOSOPHY

HARVARD UNIVERSITY

CAMBRIDGE, MASSACHUSETTS

JANUARY 2017

© 2017 Paul Julian  
All rights reserved.

## OBJECTIVITY AND INTERSUBJECTIVITY IN MORAL PHILOSOPHY

### ABSTRACT

Many people believe that morality is objective. My dissertation explores whether we have good grounds for this belief, and whether we should find it troubling if we do not. I defend negative answers to both questions. The first two chapters aim to undermine claims that we have good grounds to believe that morality is objective. The third chapter makes the case that moral normativity is essentially intersubjective, and no less respectable for that fact.

Chapter 1 poses a skeptical challenge for several promising rationales for moral objectivity. I argue that we can undermine rational confidence in these views by reflecting on the unreliability of the processes that lead us to find the views plausible, and so worth defending, in the first place.

Chapter 2 criticizes an approach that treats claims about moral objectivity as nothing more than abstract, first-order moral claims. This approach would treat the skeptical challenge as a form of substantive moral skepticism that we should reject either because it is self-contradictory or because it is implausible on first-order moral grounds. I argue that we lack good grounds to accept this approach, and so that the skeptical challenge survives.

Chapter 3 articulates and defends *Intersubjectivism*, which treats moral normativity as essentially intersubjective. I argue that we create and maintain moral normativity together by participating in relationships of mutual recognition and accountability. Intersubjectivism casts us as co-authors of morality's authority, and treats moral normativity as arising from the practical authority we grant one another to make claims within moral relationships.

## CONTENTS

<b>Introduction</b>	<b>1</b>
1. Outline of the Dissertation	1
2. Two Deviations	3
<b>1 Cutting the BS in Metaethics</b>	<b>7</b>
1. Introduction	7
2. Basic Bootstrapping	11
3. Rationales for Defending MI	14
4. What is Wrong with BS-ing, Anyway?	30
5. Conclusion	46
<b>2 On Treating Moral Objectivity as a Moral View</b>	<b>48</b>
1. Introduction	48
2. Dworkin on the Varieties of Skepticism	52
3. Dworkin on External Error Skepticism	59
4. Dworkin on Status Skepticism	72
5. Conclusion	81
<b>3 Intersubjectivism: A Theory of Moral Normativity</b>	<b>82</b>
1. Introduction	82
2. A Strategy for Reconciliation	82
3. Living in the Minds of Others	93
4. How the Drive for Respect Makes Us Persons	108
5. Defending Intersubjectivism	123
6. Conclusion	131
<b>Bibliography</b>	<b>132</b>

# INTRODUCTION

## 1. Outline of the Dissertation

Many people believe that morality is objective. This dissertation explores whether we have good grounds for this belief, and whether we should find it troubling if we do not. I defend negative answers to both questions. The first two chapters aim to undermine claims that we have good grounds to believe that morality is objective. The third chapter makes the case that moral normativity is essentially intersubjective, and no less respectable for that fact.

A person who believes that morality is objective is committed to the following view:

**Mind-Independence about Moral Normativity (MI):**

Moral normativity exists, and its existence and character are ultimately independent of the existence and character of people's moral attitudes and practices.

This position is traditionally referred to as *moral realism*. I avoid this term, however, because it misleadingly suggests that the only way moral normativity could be real is if it were independent of our thoughts about it. MI consists of two distinct claims, one existential and the other relational. The *existential claim* holds that there is a domain of facts about what there is moral reason to do or not do, or about what actions or attitudes are morally right or wrong. The *relational claim* holds that this domain of moral facts is related to subjects in a particular way—namely, objectively. To say that a fact is objective is to say that it obtains independently of the thoughts or reactions of subjects. The purported independence of moral facts would affect what it would mean to have knowledge of them: they would be there to be discovered like physical facts are, or, if you prefer, like logical and mathematical facts are.

In Chapters 1 and 2, I grant the existential claim and challenge the grounds for accepting the relational claim. Chapter 1 poses a skeptical challenge for several promising

rationales for MI. I argue that we can undermine rational confidence in MI by reflecting on the unreliability of the processes that have led us to find the view plausible, and so worth defending, in the first place. Chapter 2 criticizes an approach to defending MI that, if successful, would cause trouble for the proposed skeptical challenge. That approach treats claims about moral objectivity as nothing more than abstract first-order moral claims, and so would treat the skeptical challenge as a form of substantive moral skepticism that we should reject either because it is self-contradictory or because it is implausible on first-order moral grounds. I argue that the skeptical challenge survives this attack unscathed.

After establishing and defending the skeptical presumption against MI, I turn to articulating my positive account. We can characterize two alternatives to MI in terms of the positions they take on the existential and relational claims. **Nihilism about Moral Normativity** (Nihilism) rejects the existential claim and so also rejects the relational claim. It holds that there is no such thing as moral normativity, and so, of course, nothing of that sort capable of relating to us in one way or another. Many people seem to believe that rejecting MI commits one to endorsing Nihilism. But there is another option. **Mind-Dependence about Moral Normativity** (MD) accepts the existential claim but offers an interpretation of the relational claim that takes the existence and character of moral normativity to depend ultimately on the existence and character of people's moral attitudes and practices.

In Chapter 3, I articulate and defend **Intersubjectivism**, a form of MD that takes moral normativity to be essentially intersubjective. I argue that we create and maintain moral normativity together by participating in relationships of mutual recognition and accountability. Intersubjectivism casts us as co-authors of morality's authority, and treats moral normativity as arising from the practical authority we grant one another to make claims within moral relationships. The most pressing challenges for Intersubjectivism are to

show that it can make sense of morality's distinctive authority, its overwhelming importance, and the sense in which it is non-contingent.

## **2. Two Deviations**

Before proceeding, it may be helpful to preview two ways in which my approach throughout the dissertation deviates from much of contemporary metaethics. First, I place significant emphasis on the relevance of psychological considerations for metaethical theorizing. I pursue psychological explanations for why people might be attracted to defending MI, and I charge several MI views with operating from implausible psychological hypotheses about the practical effects of rejecting MI. Moreover, my case for Intersubjectivism appeals to the psychological dynamics of interpersonal relationships, especially our concern for others' respect. Finally, my assessments of alternative metaethical theories turn on how well they are able to help us reconcile the perspective we inhabit when engaging in moral thought with the perspective we inhabit when studying ourselves—including our dispositions to engage in moral thought—as parts of the natural world. Judging whether some theory is successful in achieving this reconciliation is, in the end, as much a psychological matter as it is a logical one.

The pride of place I grant to these psychological considerations may make my approach jarring to those whose metaethical theorizing focuses on technical issues like the semantics of moral speech or the metaphysics of supervenience. It may be even more jarring to those who see no sharp divide between metaethics and substantive ethics, and so who, following Hume, view any purely descriptive claim as more or less irrelevant to theorizing about normative matters. I will not attempt to address potential worries about this aspect of

my approach before the reader has seen it in action, though I will try to make my rationale for it clear at each step along the way.

The second respect in which my approach diverges from much of contemporary metaethics is that it focuses specifically on moral normativity, rather than on all of practical normativity or on the entire normative domain. This approach has become somewhat anachronistic. Over the past several decades, a growing number of writers recognized that whichever features of moral reasons seem to stand in need of philosophical explanation are shared by prudential reasons, as well.<sup>1</sup> It was a natural next step to see that many puzzling features of moral and prudential reasons are also shared by reasons for belief. In this way, what began as an inquiry into the nature of morality has broadened into an inquiry into the nature of normative reasons more generally. This contemporary approach has a number of significant benefits. Focusing on the general notion of a reason makes metaethics less provincial, shines a light on how thoroughly normative judgments infuse our lives, raises important issues about how different normative domains relate to one another, and confronts us with methodological questions about how general a successful metaethical theory would need to be.

Nevertheless, I depart from this contemporary approach for two reasons. First, morality is the normative domain about which people are most likely to worry that a mind-dependence theory is objectionably revisionary. Demonstrating the viability of MD is thus a crucial first step toward demonstrating that normativity in general is mind-dependent. By the same token, undermining the rationale for defending MI is a crucial first step toward undermining the rationale for defending mind-independence about normativity in general.

---

<sup>1</sup> I refer to reasons here, but feel free to insert your favored purportedly basic normative concept.

In support of this approach, I'd like to point out that metaethical theorists began by focusing on the nature and status of morality for a good reason: it is the normative domain that we most need to make sense of. Unlike disputes concerning other areas of normative discourse, disputes about the nature of morality have significant implications for how we understand our daily social lives. Much more is at stake in how we understand the wrongness of pillaging than in how we understand the prudence of brushing one's teeth after a meal. Our moral commitments shape how we conceive of and organize our lives together. They express a conception of who we are, what we care about, and how we relate to one another. It matters that we can view these moral commitments as well grounded, both for our own self-understanding and for our confidence in the claims we are disposed to make on one another. It is sensible, then, that metaethical theorists of all stripes should want to get clear about morality first and foremost.

The second reason I depart from the contemporary approach is that I am uncertain we can articulate a single, unified account of normativity capable of producing illuminating explanations of every normative domain. The divergent characters of different normative domains may call for divergent explanatory strategies. After all, what makes these domains different is that each has a different subject matter, is governed by different internal standards, features different patterns of acceptable reasoning, and so on. While I argue that moral normativity is essentially intersubjective, I am not confident that the same is true of epistemic normativity, even though I suspect that epistemic normativity is best conceived of as mind-dependent. If I were to leave out the intersubjective nature of moral normativity in the process of arguing for a more general thesis of normative mind-dependence, my case for the former would be substantially weakened. Thus, we can view my focus on moral normativity as a decision to pursue the second option in the methodological tradeoff

between trying to account for several things at a higher degree of abstraction and trying to account for one thing with a higher degree of specificity.

Of course, how successfully I accomplish this task is a separate matter. Let's now turn to the arguments.

## CUTTING THE BS IN METAETHICS

### 1. Introduction

This chapter poses a novel skeptical challenge to several promising rationales for moral objectivity. I argue that endorsing any of these rationales requires one to reason in an epistemically suspect way, and so gives one no grounds for rational confidence in the view. If my arguments succeed, advocates of moral objectivity will need to turn elsewhere for a defense.

Recall that a person who believes that morality is objective is committed to the following view:

**Mind-Independence about Moral Normativity (MI):**

Moral normativity exists, and its existence and character are ultimately independent of the existence and character of people's moral attitudes and practices.

Opponents have objected to MI on numerous metaphysical and epistemological grounds. These objections collectively give voice to the concern that objective moral facts seem especially mysterious. In response, advocates of MI have offered a battery of arguments defending their preferred version of the view. They seek to explain what kinds of things objective moral facts are (or would be), how we can (or could) discover them, how they (would) relate to other kinds of facts or entities, how they (would) connect with human interests and motivation, and so on. However, one kind of argument is conspicuously rare. This argument would explain why we should conceive of morality as objective to begin with. Granted that we *can* offer an account on which morality is objective, what reason do we have to think that it *is* objective, and that an account of its objectivity is worth defending?

Consider an analogy. We know a lot about centaurs.<sup>2</sup> We know that they are composed of a human head, arms, and torso atop the body and legs of a horse. We know that they possess enviable strength and a stately, upright posture. We can debate whether they were brought into existence by Ixion and Nephele (a cloud in the shape of Hera), or instead by Centaurus and a Magnesian mare. Now, what kind of evidence would you need for it to seem reasonable that centaurs are not just mythological creatures we have created to spice up our stories, but that centaurs exist (or have existed) mind-independently, populating the physical world alongside familiar animals like leopards and badgers? Perhaps you'd need to see and interact with one while being confident you weren't losing your mind, or maybe you'd need to hear plausible testimony about them from reputable biologists or paleontologists.

We also know a lot about morality. We know, for instance, that it is wrong to trip a stranger in order to embarrass her in public. We can debate whether this is because it makes the stranger suffer, or because it treats her with a lack of respect, or because it expresses something bad about one's own character. Now, what kind of evidence would you need for it to seem reasonable that morality is not just a distinctive and pervasive expression of human concerns, which we have made clearer, more consistent, and more systematic over time, but that morality is objective, like certain physical or mathematical facts are? Well, unlike centaurs, objective moral facts are not the kinds of things you could observe in the wild, or that you could be wooed or attacked by, or (arguably) that you could justifiably believe in solely on the basis of testimony from self-proclaimed experts. From whence, then, does the commitment to defending MI originate, and how good are the reasons supporting it?

---

<sup>2</sup> Actually, I do not know a lot about centaurs. What follows is from Wikipedia.

In this chapter, I take up these last two questions. I aim to show that several of the most promising and widespread rationales for defending MI rely on a suspect form of reasoning called *bootstrapping*, or, more concisely and pejoratively, *BS-ing*. I then argue that one should lack confidence in a conclusion reached through BS-ing if one has evidence that this BS-ing is unreliable. I argue that we have such evidence in the cases under discussion, and thereby defend the skeptical claim that we lack grounds for rational confidence in MI on the basis of these widespread rationales.

Let me now explain why this skeptical approach is novel. Traditional moral skeptics come in many stripes and defend epistemological claims of differing strengths. Some argue that if there were such things as objective moral facts, we could only access them by employing a unique and strange cognitive faculty, which we have no reason to believe we possess.<sup>3</sup> Others argue that because evolution has so thoroughly shaped our evaluative tendencies—in ways that are adaptive rather than in ways that track an independent moral truth—we should admit that even if objective moral facts did exist, we would almost certainly be clueless about their content.<sup>4</sup> Still others argue that the best explanation of widespread and persistent moral disagreement among people who seem to be epistemic peers is that different moral beliefs reflect different ways of life, say, or different outcomes of competitions for social power, rather than that there is an objective moral reality that some “perceive” more successfully than do others.<sup>5</sup>

While these varieties of moral skepticism are quite different in their details, we can think of them as posing the same general tri-lemma: either abandon MI, abandon the claim

---

<sup>3</sup> See J.L. Mackie, *Ethics: Inventing Right and Wrong* (New York: Penguin, 1977), Ch. 1.

<sup>4</sup> See Sharon Street, “A Darwinian Dilemma for Realist Theories of Value,” *Philosophical Studies* 127 (2006), pp. 109-66.

<sup>5</sup> See, e.g., Mackie (1977); Gilbert Harman, *The Nature of Morality: An Introduction* (Oxford: Oxford University Press, 1977); Friedrich Nietzsche, *On the Genealogy of Morality*, ed. Maudemarie Clark and Alan J. Swensen (Indianapolis: Hackett Publishing, 1998).

that we have moral knowledge, or provide some account of how knowledge of objective moral facts is possible. Defenders of MI have been happy to try their hands at the third prong. And so the debate has continued in terms of whether and how MI could be consistent with our having moral knowledge, or at least justified moral belief.

The skeptical argument I construct in this chapter is importantly different from these more familiar approaches, as it poses the challenge one step earlier in the dialectic. It queries directly what reason we have for conceiving of morality as mind-independent. Thus, we might think of the argument as raising a challenge not within moral epistemology, but rather within metaethical epistemology. Rather than questioning our justification for holding first-order moral beliefs, the argument questions our basis for holding the second-order belief that our moral beliefs are objectively true or false. Some philosophers link the two issues by holding that moral discourse presupposes the truth of MI, such that, e.g., “is morally wrong” entails “is objectively morally wrong.”<sup>6</sup> This dissertation is largely an attempt to show why this linking move is a mistake.<sup>7</sup>

The chapter is organized as follows. In §2, I illustrate how BS-ing works and explain why it is epistemically troubling. In §3, I show that several rationales for defending MI are guilty of BS-ing. I then argue that employing a BS-ing rationale for MI gives one no grounds for rational confidence in the view. Finally, in §4, I consider objections that BS-ing in favor

---

<sup>6</sup> This is true not only of many MI theorists, but also of many error theorists. For example, see Richard Joyce’s insistence that moral concepts have “practical clout,” which he understands as committing one to something like MI. See *The Evolution of Morality* (Cambridge, MA: MIT Press, 2006), Chapters 2 and 6.

<sup>7</sup> If one holds the linking view, then one will think the argument of this chapter attempts to undermine our justification for all of our first-order moral beliefs, as well, which would make the argument seem more ambitious and revisionary than it is. But if the argument of this chapter is plausible, it gives defenders of the linking view a good reason to rethink it, or else risk throwing the baby out with the bathwater.

of MI is not epistemically problematic, and argue that these objections fall in the face of two undercutting defeaters.

## 2. Basic Bootstrapping

Imagine that you seek out the new village shaman. You have suffered through several poor harvests in a row, and you think she may be able to help. Sure enough, the shaman is ready with advice. She tells you that the way to ensure the success of your next harvest is to placate the gods. And the way to placate the gods, she says, is to sacrifice a goat. This approach seems promising. After all, you have spent the past few years ignoring the gods and focusing on earthly details like tilling, seeding, and watering.

You do not have many goats. But you know that if your crops fail, you will have to eat the goats anyway, and even that might not sustain you through the winter. So, you sacrifice a goat. When you see the shaman again the following week, she clarifies her edict: the gods are not placated by a solitary act so much as by consistent devotion. This is bad news. It means you will need to sacrifice another goat. Begrudgingly, you do. The third week brings worse news: many of your fellow villagers have also begun to follow the shaman's advice. Since the gods can only have so many favorites, you must now distinguish yourself by sacrificing two goats in each ritual.

You'll soon run out of goats at this rate, so you decide to consider more carefully whether the shaman is a trustworthy guide to matters of crop cultivation. Suppose you reason like this:

### **A CROPPY ARGUMENT (CROP):**

- (1) The shaman has repeatedly told me that in order to ensure a successful harvest, one must sacrifice goats to placate the gods.
- (2) Moreover, it is true that in order to ensure a successful harvest, one must sacrifice goats to placate the gods.

- (3) Thus, the shaman has repeatedly been accurate (and never inaccurate) in telling me what will ensure a successful harvest.
- (4) Thus, the shaman is a reliable source of beliefs about what will ensure a successful harvest.
- (5) If a belief source is reliable, then one is justified in trusting it, i.e., in believing its outputs.
- (6) Thus, I am justified in trusting the shaman's advice about what will ensure a successful harvest.<sup>8</sup>

On this basis, you decide to maintain your trust in the shaman. You thereby doom your goats for no good reason.

CROP is not a very good argument. It might be valid if we added some relatively benign premises, but I'd like to set aside concerns with the details<sup>9</sup> and focus instead on what is wrong with your *use* of the argument. The problem, as you may have guessed, is that your use of CROP is circular. The argument itself is not logically circular, since the conclusion does not appear as a premise in the chain of reasoning. Rather, your use of the argument is *epistemically circular*: you presuppose the conclusion in the process of arguing for it. You presume (4) in accepting (2), yet you purport to use (2) to argue for (4). By hypothesis, you did not accept (2) before speaking with the shaman, and you have no independent reasons to accept (2), perhaps aside from the idea that the shaman has privileged access to supernatural or agricultural facts—a claim for which you so far have no evidence. Yet, in believing (2), you treat the shaman as reliable, even though the shaman's reliability is the very issue in question.

While CROP is a track-record argument, not every instance of epistemic circularity involves using arguments of that form. What is essential for some piece of reasoning to be

---

<sup>8</sup> This way of framing the argument is inspired by Michael P. Lynch and Paul Silva, Jr., "Why Worry About Epistemic Circularity," *Journal of Philosophical Research: Supplement—Papers in Honor of William P. Alston* (May 2016).

<sup>9</sup> For instance, some might resist the move from (3) to (4) on the basis that one would need more than a 3-0 track record of success before labeling a belief source reliable. Others might offer substantive objections to (2) on a number of different grounds, or reject (5) because they do not accept reliabilism about epistemic justification.

guilty of epistemic circularity is that one appeal to or rely upon some belief source in the very attempt to vindicate the epistemic standing of that source. Many philosophers agree that there is something fishy about epistemically circular reasoning, but there is significant disagreement about what, if anything, is ultimately wrong with it. On my view, the problem with epistemically circular reasoning is that it is rationally inert. It is incapable of accomplishing the goal of reasoning, which is to help the reasoner decide what to believe. Since your acceptance of needed premises depends upon your already treating the conclusion as true, reasoning in an epistemically circular manner cannot help you understand whether you rationally ought to believe that conclusion.<sup>10</sup> Thus, if your only argument in favor of a conclusion requires you to reason circularly, you lack grounds for rational confidence in that conclusion.

From here on, I shall refer to the use of epistemically circular reasoning by its more evocative synonym—*bootstrapping*, or *BS-ing*—since the latter term more vividly expresses the impropriety of attempting to vindicate a belief source by treating its outputs as “testimony” or evidence of the source’s trustworthiness. The BS-er tries to leverage a pre-existing commitment to the truth of some beliefs into showing that one is justified in relying upon the beliefs’ source. As in the case of the person who tries to pull herself up by her own bootstraps, the BS-er seeks leverage where none is to be found. This approach leaves the BS-er exactly where she began, the epistemic equivalent of sitting on the ground.

---

<sup>10</sup> This claim may seem unfairly to favor or presuppose internalism about justification or other epistemic merits. I will consider externalist responses in §4.

### 3. Rationales for Defending MI

#### 3.1 Interpreting Moral Experience

So, what does all this have to do with defending moral objectivity? I now want to consider how a commitment to defending MI might originate. Then, I will argue that the rationale underlying that commitment is likely to involve at least one of several forms of BS-ing.

The appeal of defending MI most plausibly stems from reflecting on moral experience itself. When we think about what to do, when we make a claim on another, when we feel guilty, or when we levy blame for some transgression, it may seem as if we are appealing to standards that exert normative sway over us independently of our granting them that power. Indeed, sometimes we consider ourselves bound by these standards even while wishing we were free of them entirely. Reflecting on moral experience can, in this way, make compelling the idea that the moral domain is marked by a kind of mind-independence.

The case for this idea seems clearest when one considers moral beliefs that one takes to be categorical and in which one is very confident. For instance, suppose you believe that it would be wrong to pillage a neighboring village for sport. Such pillaging would cause enormous avoidable suffering, would treat its victims as mere means to one's demented satisfaction, and would express some of the worst characteristics of which humans are capable. You can think of no circumstances in which it would be permissible, and no consideration that would make you change your mind about its wrongness. If you encountered a marauder who cared not a whit about the above considerations, you would think she was making a grave moral mistake. She would be failing to recognize moral constraints that applied to her. It may seem natural, then, to conclude that pillaging for sport is not just morally wrong, but *objectively* morally wrong.

You might draw a similar conclusion just by reflecting on what it's like to reconsider or revise moral beliefs. Imagine that Helen is your village honcho. You once believed that because Helen is a direct descendent of the village founder, she alone has the right to decide the rules. But suppose you have come to question this belief on the basis of a different and, you think, more plausible set of beliefs. Perhaps you think that it shouldn't be arbitrary who gets to decide the rules, and that it is arbitrary that Helen happens to have been born into her position. These latter beliefs convince you that your original belief about Helen's favored status was mistaken, that although it once seemed to you to be correct, you now see the ways in which it neglected crucial moral features of the case. And what makes your new beliefs correct seems to be their content, not the fact that you take them to be correct. Reflecting on this process of scrutinizing and revising your moral beliefs may reinforce your overall sense that you are dealing with matters that are independent of your attitudes about them. In changing your moral beliefs, you apparently acknowledge, and perhaps even presuppose, that the standards for moral correctness outrun the beliefs themselves.

These observations in the pillaging and honcho cases both occur within the realm of moral phenomenology—the “what it's like” of moral experience—and this seems the most natural place to begin our theorizing. After all, whenever we set out to give a philosophical account of some domain, we have to bring with us an everyday conception of what is to be explained, and morality is first and foremost a practical, lived subject. Let's call the phenomenological conception advocated in these two cases the *objectivity interpretation of moral experience*. We should consider at least two issues in assessing this conception. The first issue is broadly interpretive, and concerns whether MI is the best, or only, account that comports with the phenomenological data. The second issue is broadly methodological, and concerns

the extent to which we should treat a phenomenological conception as the bedrock upon which our accounts are built, as opposed to merely its raw material.

Let's begin with the interpretive issue. Do the phenomenological data—the way things seem when we hold, question, or revise a moral belief—suggest, or even necessitate, a commitment to moral objectivity? In the village honcho case, your process of belief revision consisted of critiquing one moral belief in light of another. You considered whether Helen's genealogical status grants her the authority to decide the rules and decided instead that genealogical status is irrelevant for political authority. But what evidence do you have that these moral claims are made true or false in virtue of their corresponding to *objective* moral facts? After all, it seems that we could explain your deliberative process more parsimoniously just by appealing to the fact that you weighed a commitment to one moral claim against a commitment to another with which the first is inconsistent, and decided in favor of the latter.

Nevertheless, we would still need to account for the felt sense that the normative force of these claims is somehow *independent* of us. And, as I mentioned at the outset, there is a strong conceptual link between independence and objectivity. In §4.3, I will introduce an alternative interpretation of this link that attempts to account for our phenomenological data by appealing to intersubjective considerations, rather than to objective ones.

### **3.2 Phenomenal Access to MI**

Suppose, for the time being, that you decide to maintain the objectivity interpretation of moral experience. Let's now turn to the second, methodological issue: is there a legitimate rationale for defending MI on the basis of this interpretation of the moral phenomenology? Here is a simple but implausible view:

**Phenomenal Access to MI** (“Phenomenal Access”)

We can justifiably draw a metaethical conclusion from phenomenological premises, so we are justified in believing MI on the basis of (the objectivity interpretation of) moral experience.

It would be phenomenal if Phenomenal Access were true. It would also be surprising, since in no other domain do we license claims about objectivity on the basis of phenomenological claims. Thus, a defender of the view would need to give us especially good reasons to think that reflecting on moral experience is a reliable guide to metaethical truth.

Just to clarify what would be required, let’s say that “moral experience” involves considering substantive moral propositions, as well as forming, expressing, questioning, revising, or being motivated to act on one’s moral beliefs or commitments.<sup>11</sup> Furthermore, let’s say that “ $P_1$ - $P_n$ ” picks out a set of substantive moral propositions in which a given person is confident. For instance, most readers would take  $P_1$ - $P_n$  to include the propositions that pillaging for sport is wrong and that a person’s genealogy is irrelevant to their authority over others. Finally, let’s assume that  $P_1$ - $P_n$  is sufficiently large that it seems reasonable to make generalizations about morality as a whole on the basis of one’s beliefs about the various propositions contained in the set.

I contend that if you were to defend Phenomenal Access, your reasoning would have to look something like this:

**ARGUMENT FOR PHENOMENAL ACCESS** (“PHENOMENAL”)

- (7) My moral experience presents  $P_1$ - $P_n$  as true or false in a mind-independent way. [objectivity interpretation of moral experience]
- (8)  $P_1$ - $P_n$  are true or false in a mind-independent way. [commitment to MI]

---

<sup>11</sup> This characterization of moral experience might seem both insufficiently practical and insufficiently social, especially given the positive view I go on to defend. In particular, one might think that the primary activities that constitute our moral experience involve *actually interacting* with others by making and responding to moral claims. But characterizing moral experience in terms of these interactions would set the bar too high. We are constantly thinking in moral terms—considering whether we or others are acting permissibly, selfishly, kindly, and so on—and many of these thoughts never spur us to act or interact, though of course they organize our attitudes in ways that guide us when we do act and interact.

- (9) Thus, my moral experience is consistently accurate in indicating that MI is true.
- (10) Thus, my moral experience is a reliable source of beliefs about whether MI is true.
- (11) If a belief source is reliable, then one is justified in trusting it, i.e., in believing its outputs.
- (12) Thus, I am justified in believing MI on the basis of my moral experience.

Using PHENOMENAL to reach a conclusion in favor of MI involves BS-ing in an identical way to using CROP to reach a conclusion in favor of sacrificing your goats. Here, you presume (10) in accepting (8), since you must already take your moral experience to be a reliable metaethical guide in order to take yourself to have reason to claim that your moral beliefs are true in a mind-independent way. While in CROP you presume the shaman's trustworthiness when that is the very issue in question, in PHENOMENAL you presume the metaethical trustworthiness of your moral experience when that is the very issue in question. Put differently, you BS your way to the conclusion by taking the outputs of (the objectivity interpretation of) moral experience—the beliefs that  $P_1$ - $P_n$  are true or false in a mind-independent way—as evidence that your interpretation of moral experience is a reliable guide to whether morality is, indeed, objective. If we take your use of CROP to be problematic in virtue of its BS-ing, we ought to take the same stance toward your use of PHENOMENAL.

It is important to appreciate the generality of PHENOMENAL and of the worries about it. Suppose one is a rationalist and objects to all this talk about gaining insight into metaethical truths by reflecting on one's moral experience, since the focus on experience sounds too *a posteriori*. The way we access metaethical truths, the rationalist would claim, is through the use of rational intuition, not through experience. But this objection misunderstands PHENOMENAL, which is neutral with respect to the nature of the moral

faculties that make one's moral experience possible.<sup>12</sup> Moreover, it is essential neither that one comes to believe  $P_1$ - $P_n$  over a period of time, nor that  $P_1$ - $P_n$  contains more than one believed moral proposition.<sup>13</sup> All that matters is that one call upon a belief in attempting to demonstrate the reliability of the source of that belief. In sum, the objection here targets the general process of checking a belief source against itself, by taking that source's outputs as evidence of its reliability, and in the absence of any independent, corroborating support.

In §3.4, I will consider an alternative way one might defend MI by appealing to rational intuition. But first, I use the next section to consider *why* someone might be tempted to BS. My case so far may seem to be weakened by the fact that using CROP or PHENOMENAL to reach their respective conclusions seems so obviously wrongheaded. It may seem implausible, then, that otherwise rational people would BS in this way. I next offer an approach to understanding the motivation behind BS-ing that, if accurate, would strengthen my case that BS-ing often underlies the rationale for defending MI.

### 3.3 Practical Commitment and Practical BS-ing

I mentioned in §1 that it is surprisingly rare to find a proponent of MI who explicates the foundational rationale for defending the view. One notable exception is David Enoch. He offers the following “confession”:

I suspect that as a psychological matter, I hold the metaethical view I in fact hold not because of highly abstract [philosophical] arguments...Like many other realists (I suspect), I pretheoretically feel that nothing short of a fairly strong metaethical realism will vindicate our taking morality—or perhaps normativity more generally—seriously...Metaethical positions that are not objectivist in some important intuitive sense have—in the context of interpersonal disagreement and conflict—implications

---

<sup>12</sup> Recall that moral experience, as I characterize it, is constituted by considering substantive moral propositions, as well as forming, expressing, questioning, revising, or being motivated to act on one's moral beliefs or commitments.

<sup>13</sup> On the latter point: of course, most people would find a track-record argument for the reliability of a belief source more compelling the more examples we have of its accuracy.

that are objectionable on first order, moral grounds, and should therefore be rejected.<sup>14</sup>

The “first order, moral grounds” Enoch has in mind highlight a crucial difference between how we ought to react to moral disagreements versus how we ought to react to disagreements grounded in mere preferences.<sup>15</sup> When one disagrees morally with another person, it would be morally wrong not to stand one’s ground, e.g., by agreeing to act in accordance with the other’s view, or by simply suspending judgment on the issue in question. Not standing one’s ground is tantamount to not “taking morality...seriously.”

If one’s conflict with another person is at the level of mere preference, however, Enoch claims that it would be morally wrong *to* stand one’s ground, because doing so would express an objectionable form of partiality by treating one’s own preferences as more important than the other person’s. Enoch argues that MD theories are unable to explain why we would be justified in treating moral claims as more than mere preferences, and so fail to explain why it would be morally wrong to back down in a moral disagreement. Thus, he concludes that we have substantive moral reason to believe that morality is objective: if we were to treat morality as non-objective, we would fail to take morality seriously (and, of course, we are morally required to take morality seriously).<sup>16</sup>

While Enoch’s argument is idiosyncratic, I take it to be voicing a version of the familiar worry that if MI were false, we would have no justifiable basis for maintaining our moral beliefs in the face of disagreement. This is a frightening prospect. After all, our moral

---

<sup>14</sup> David Enoch, “How Objectivity Matters,” *Oxford Studies in Metaethics*, Vol. 5 (2010), p. 111.

<sup>15</sup> Enoch focuses on such disagreements in contexts of joint action, but this distinction arises in other practical contexts, as well.

<sup>16</sup> It is unclear whether Enoch interprets this requirement to take morality seriously as an *additional* moral factor over and above whatever moral reasons apply to us in a given case, or rather as a type of *generalization* about what is entailed by the fact that those reasons apply to us. I suspect the latter interpretation is more in line with his view, though it does seem odd to claim, essentially, that morality requires you to do what morality requires you to do (and not, e.g., to abandon doing what morality requires you to do just because someone else disagrees).

beliefs would then carry the same practical weight our preferences do, and there is no overarching standard for adjudicating between preferences in certain kinds of conflicts. I might prefer chocolate, you might prefer vanilla, and neither of us would be wrong. By contrast, we think that if a marauder were trying to lop off my head, she *would* be acting wrongly. My demanding that she lower the axe would not be merely a boorish insistence that my preference be satisfied and hers frustrated, as it would if I demanded that we order the chocolate ice cream to share just because that's the flavor I'd rather have. I would be making a distinctively moral demand, the normative weight of which would trump that of the marauder's preferences. While preferences are *merely* subjective, Enoch would say, moral claims seem to involve appeals to objectivity—to standards that are not subject to a person's whims—which would give us a justifiable basis for not backing down in a moral conflict.

I think Enoch's argument is unsatisfactory for several reasons, not least of which because I think that at least one MD theory—as it happens, the one I defend in this dissertation—shows how we can treat moral claims as more than mere preferences. But let's begin with a broader point. The fact that Enoch calls his statement a confession, rather than a *profession*, say, or an explanation, suggests that it reveals something that he thinks, at some level, it would be better to keep concealed. Why might he feel inclined to keep his motivations for defending MI a secret? Perhaps it is because he rightly predicts that those motivations will seem epistemically suspect once revealed. Enoch is admitting that his decision to defend MI is based on his commitment to a substantive moral claim. More precisely, he is defending the claim that morality is objective on the basis that, if morality were not objective, then a substantive moral claim that he thinks is true would actually be false. But, he asserts, this substantive belief is true, so morality is objective. There is a certain

kind of circularity in Enoch’s reasoning, but it seems importantly different from the epistemic circularity that plagues uses of CROP and PHENOMENAL.

Let’s take a closer look. Let ‘P’ be the substantive moral claim that “when one is in a moral disagreement with another person, it would be morally wrong not to stand one’s ground.” Enoch’s argument runs as follows:

**ENOCH’S PRACTICAL COMMITMENT (“COMMITMENT”)**

- (13) We are justified in taking morality seriously if and only if MI is true.
- (14) If MI is false, then P is false.
- (15) But P is true.
- (16) So, MI is true.
- (17) So, we are justified in taking morality seriously.

COMMITMENT appears to be a valid argument. While we could raise objections to several of its premises, let’s focus, as before, on Enoch’s *use* of the argument to reach its conclusions. Enoch appeals to the truth of P in order to conclude that MI is true. However, he has characterized P in a way that *presupposes* the truth of MI. According to Enoch, to say that it would be morally wrong not to stand one’s ground in a moral disagreement *just is* to take morality seriously, or at least part of it. Moreover, (13) holds that only the truth of MI can justify us in taking morality seriously. If one accepts (13), then, in taking oneself to have grounds for accepting (15), one must already presuppose (16). In other words, Enoch uses COMMITMENT to argue in a circle.

The type of circularity at work here is distinctive. As Enoch’s “confession” suggests, it is his *practical commitment* to being able to “take morality seriously” that runs the show. Enoch leverages his commitment to treating substantive moral claims as true—expressed by his commitment to treating P as true—plus a claim that MI would need to be true in order for those substantive claims to be true, in order to argue for the truth of MI. Let’s call this variety of the circular maneuver *practical bootstrapping*. Practical BS-ing and regular, epistemic BS-ing are not mutually exclusive. In fact, I want to suggest that we can understand at least

some epistemic BS-ing as being motivated by practical BS-ing. In any domain where a skeptical worry threatens to pull the rug out from under our feet, our commitment to being able to treat everyday beliefs as justified may motivate us to practically BS.

I ended the previous section by asking why an otherwise rational person would be attracted to BS-ing in favor of MI. Here is my proposed answer: if one is already practically committed to treating MI as true, then it is easier to understand why one would presuppose MI in the process of arguing for it. We should not be surprised that practical BS-ing leads to unjustified conclusions in the moral domain. After all, consider how the move would work in other domains. Following Enoch, suppose someone tells you,

Like many other stone-worshippers (I suspect), I pretheoretically feel that nothing short of a fairly strong account of the divinity of stones will vindicate our taking nature seriously. Natural accounts that are not stone pious in some important intuitive sense have—in the context of our everyday respect for our natural surroundings—implications that are objectionable on first order, religious grounds, and should therefore be rejected.

Or consider this alternative:

Like many other lovers (I suspect), I pretheoretically feel that nothing short of a fairly strong account of soul mates will vindicate our taking romantic love seriously. Accounts of romantic love that are not soul-mate-centric in some important intuitive sense have—in the context of our everyday reliance on the universe's plan for our happiness—implications that are objectionable on first order, love-related grounds, and should therefore be rejected.

As I hope these analogies make clear, practical BS-ing gives rise to an especially egregious form of circularity. Even if our first-order experience in a domain is shaped in important ways by the second-order theory we accept, we cannot argue for the truth of that second-order theory on the grounds that doing so is necessary for preserving or vindicating our first-order experience. To do so is to let a kind of wishful thinking infect our theorizing.

In this section, I have aimed to show that there is a distinctive species of BS-ing, practical BS-ing, that is not only interesting in its own right, but that also helps to explain the

motivation behind BS-ing more generally. In the next section, I consider a way of appealing to moral experience to justify defending MI that aims to avoid the BS-ing generated by appealing to PHENOMENAL.

### 3.4 Phenomenal Conservatism

Recall from §3.2 that Phenomenal Access holds that we are justified in believing MI on the basis of (the objectivity interpretation of) moral experience. The problem was that using PHENOMENAL to defend Phenomenal Access involves BS-ing. However, several philosophers have argued that we can defend (something like) Phenomenal Access by way of a different rationale. Michael Huemer defends

**Phenomenal Conservatism:**

If it seems to S that P, then, in the absence of defeaters, S thereby has at least some justification for believing that P.<sup>17</sup>

The basic idea here is that it's reasonable to assume that things are as they seem unless we have some reason to doubt that they are. For instance, if it seems to you that a band of marauders is approaching your village—you see their snarling faces, you smell their torches, you hear their horses neighing, you feel the ground beneath you shaking—then you have at least *prima facie* justification for believing that marauders *are* approaching. What would count as a defeater for this belief being justified? Imagine you had taken mescaline this morning, or that you knew your fellow villagers routinely took delight in engineering elaborate ways to terrify you.

The phenomenal appearances in this example are primarily perceptual: they involve seeing, smelling, hearing, or feeling the marauders approaching. So how might Phenomenal

---

<sup>17</sup> Michael Huemer, “Compassionate Phenomenal Conservatism,” *Philosophy and Phenomenological Research* 74 (2007), p. 30.

Conservatism help a defender of MI? Huemer argues that the relevant form of appearance in this case is the moral intuition:

An intuition that P is a state of its seeming to one that P that is not dependent on inference from other beliefs and that results from thinking about P, as opposed to perceiving, remembering, or introspecting. A [moral] intuition is an intuition whose content is [a moral] proposition.<sup>18</sup>

So, for instance, we presumably share the moral intuition that pillaging for sport is wrong.

When we consider what pillaging for sport involves, its moral wrongness appears obvious to us.<sup>19</sup> This intuition is a kind of *intellectual appearance*, not unlike our intuition that  $5 - 3 = 2$ .

Suppose we grant Huemer this claim. Next, we should want to know on what basis we should believe that pillaging for sport is not just morally wrong, but that its moral wrongness is *objective*, and thus that MI is true.

Huemer argues that we should understand Phenomenal Conservatism as a form of direct realism. Here's how the view treats sensory experience:

...[T]he primary function of sensory experience is to partly *constitute* our awareness of external things, rather than to be an intermediary *object* of awareness...[W]e are first of all aware of *things*—that is, external things. Then we reflect on our awareness of those things, whereupon we notice (or perhaps infer) that such awareness involves our having internal states that somehow represent external things. These internal states should not be allowed to supplant the real objects in our philosophy; their central function is that of vehicles of the awareness of external things.<sup>20</sup>

The analogue for moral experience would hold that the moral facts we intuit are “out there,” in some non-spatial sense, independent of our intuiting them. How do we know this?

Huemer defends direct realism by treating the objectivity interpretation of moral experience as *itself* an intuition. So, first, we have substantive moral intuitions. Then, when we reflect on our experience of forming substantive moral intuitions, we form the intuition that this

---

<sup>18</sup> Michael Huemer, *Ethical Intuitionism* (New York: Palgrave Macmillan, 2005), p. 239.

<sup>19</sup> Or perhaps we infer from some more basic moral propositions that pillaging for sport is wrong. If so, then *those* propositions would then be the ones for which we have the corresponding intuitions.

<sup>20</sup> *Ibid.*, p. 243.

process involved representing objective moral facts. On this approach, if it seems as if moral facts obtain objectively, then we have *prima facie* justification to believe that they do obtain objectively, and thus that MI is true.

Phenomenal Conservatism may seem too good to be true, since it licenses *prima facie* justified belief so easily. After all, taking stock of the way things appear seems to be only the first step in determining what the world is like. Don't we then need some sort of independent evidence, above and beyond our faith or confidence in the way things appear, in order to vindicate those appearances? Moreover, notice that the pattern of reasoning in defense of direct realism is eerily similar to that in PHENOMENAL, which raises the question whether Phenomenal Conservatism is just a dressed-up version of the same view. Just how different is it to claim, as Phenomenal Access does, that one's phenomenological interpretation of moral experience is a reliable basis for believing MI, as opposed to claiming, as Phenomenal Conservatism does, that the appearance that morality is objective provides *prima facie* justification for believing MI?

There are two major differences between these approaches. The first concerns the level of justificatory support being posited. The *prima facie* justification posited by Phenomenal Conservatism is a weaker status, and leaves it open whether one's justification for believing MI based on appearances will, in the end, be undermined by other considerations (Huemer obviously thinks not, but below I argue otherwise). The second major difference is that the proponent of Phenomenal Conservatism offers a powerful general rationale for granting this *prima facie* justificatory status to beliefs based on appearances.

Huemer argues that while it may initially seem reasonable to require independent support for such beliefs, taking this path opens the door to global skepticism, and so is self-defeating. He writes,

What happens if we apply the principle generally: 'We need positive reasons for trusting appearances'? Then we need positive reasons for trusting sense perception, memory, introspection, even reason itself. The result is global skepticism. Nothing can be accepted until we first give a positive reason for trusting that kind of belief. But we cannot give such a reason without relying on sense perception, memory, introspection, reason—or in general, on *some* source. Hence, we shall never be able to trust anything. Of course, this means we also could not trust the reasoning of this paragraph.<sup>21</sup>

Huemer argues that all of our beliefs are based on how things appear to us. If we were to demand to know how we can trust those appearances, we would be requesting evidence that basing our beliefs on appearances leads us to represent the world accurately. What would count as the relevant evidence? We would presumably need to show that our belief-forming mechanisms are trustworthy. How could we show this? The most natural way would be by pointing to instances in which basing our beliefs on appearances lead us to represent the world accurately. But if we took this approach, we would be presupposing the trustworthiness of the very belief sources we were seeking to vindicate. This approach would thus implicate us in BS-ing. Moreover, any alternative approach would, by necessity, require us to appeal to an appearance of some other sort—since all thinking is based on appearances—merely pushing the problem back one level and courting the same circularity with regard to a different belief source. Thus, Huemer concludes, we have a general rationale for taking ourselves to be *prima facie* justified in forming beliefs based on appearances: not

---

<sup>21</sup> Ibid., p. 244.

doing so leads to global skepticism. He argues that it makes more sense to treat trust in appearances as a kind of “rational default position.”<sup>22</sup>

This line of reasoning suggests two possible ways to defend belief in MI. First, of course, we could accept Phenomenal Conservatism, in which case our work would be finished. But, second, even if we ended up rejecting Phenomenal Conservatism, we find in Huemer’s reasoning a possible general defense of BS-ing. Notice that Huemer pitches his view as a way to avoid a version of what Chisholm calls the “problem of the criterion,” one of the oldest and most difficult puzzles in epistemology.<sup>23</sup> In this version of the problem, it seems we cannot justify our appearance-based beliefs unless we have some criterion that helps us determine which of those beliefs is justified. However—and here is where the puzzle emerges—we cannot justifiably believe that some criterion is trustworthy unless we can show that it picks out only the *justified* appearance-based beliefs. And, so, we are back to square one, unable to pick out the correct criterion because we are unable to pick out the justified beliefs, and vice versa. Versions of this puzzle include the Cartesian circle, Hume’s problem of induction, and, most relevant for our purposes, worries about conclusions reached through BS-ing. But if these latter worries are merely one instance of a much more general worry about how to vindicate some capacity without relying on that capacity, then BS-ing in defense of MI is just a form of this very general problem, such that one could mount a companions-in-guilt argument holding that there is nothing *in particular* wrong with defending MI in this way.

I will reply to both of these defenses in §4.3. Rather than demand a positive reason to believe that our moral intuitions are reliable guides to metaethics, I will argue that we have

---

<sup>22</sup> Michael Huemer, “Phenomenal Conservatism,” *Internet Encyclopedia of Philosophy* (<http://www.iep.utm.edu/phen-con/print#SH3d>).

<sup>23</sup> Roderick M. Chisholm, *The Foundations of Knowing* (Minneapolis: The University of Minnesota Press, 1982). Discussion of the problem goes back at least to Sextus Empiricus.

independent reason *to doubt* that our moral intuitions are reliable in this way. If successful, this argument will serve as an undercutting defeater for belief in MI on the basis of Phenomenal Conservatism. I will call on the same considerations to argue that BS-ing in favor of MI is more like BS-ing in CROP than it is like BS-ing in attempting to vindicate basic capacities like introspection or sense experience. Thus, I will argue that our ability to distinguish between various types of BS-ing allows us block a companions-in-guilt defense of BS-ing in favor of MI.

### 3.5 Two Other Rationales for Defending MI

Before we proceed to that discussion, I want to mention briefly and then set aside two other possible rationales for defending MI. The first rationale holds that there is no satisfactory version of MD on offer, and so, in light of the objectivity interpretation of moral experience discussed above, it is reasonable to attempt to defend a version of MI. Let's call this rationale "Conditional Support," since its support for MI depends on the judgment that the prospects are bleak for articulating a successful form of MD. Conditional Support is a reasonable view. I suspect that many who endorse MI based on how it fares in wide reflective equilibrium can be understood as tacitly appealing to Conditional Support. Nevertheless, we will have reason to abandon Conditional Support if Chapter 3 of this dissertation succeeds in its articulation and defense of Intersubjectivism as a novel version of MD.

The second rationale has become increasingly popular over the past two decades. Let's call it

**Substantive Stance:**

Any claim that is relevant to how we understand the nature and status of morality is itself a substantive, moral stance.

According to this view, we must treat both MI and MD as general moral claims, rather than as metaethical claims. Several prominent proponents argue that once we accept Substantive Stance, it becomes obvious that we should also accept MI. I devote Chapter 2 to arguing against this view.

#### **4. What is Wrong with BS-ing, Anyway?**

Our next step is to consider objections to my claim that the BS-ing we've identified is a genuine problem. I will consider the two objections that seem most promising, and argue that they both fail. First, I will consider the objection that there is nothing wrong with BS-ing as such, since justification is best understood in an externalist fashion. Second, I will consider the objection that we should not be troubled by BS-ing in favor of MI because we cannot vindicate any basic belief source without BS-ing. I will argue that we have undercutting defeaters for both of these objections, as well as for the approach that appeals to Phenomenal Conservatism to defend belief in MI.

##### **4.1 Solace in Externalism?**

In §2, I claimed that the problem with using BS-ing reasoning is that it cannot help you understand whether you ought to believe its conclusion, since your acceptance of the needed premises depends upon your already treating the conclusion as true. I claimed that this fact prevents you from having *rational confidence* in a conclusion reached (only) through BS-ing, making your reasoning rationally inert. Rational confidence, as I understand the concept, is a measure of credence one should have in a proposition in light of one's (actual or potential) awareness of the rational support for it.

I have so far avoided making claims about how BS-ing affects one's *justification* for believing some conclusion, because I do not want to unfairly presuppose internalism about justification. But one might naturally wonder how the notion of rational confidence relates to an internalist conception of justification, since both assess epistemic status in terms of factors that are in some sense internal to an agent. In this section, I consider the externalist view that one can use a BS-ing argument to gain justification for believing a conclusion. I employ the concept of rational confidence to help explain the epistemic defect of BS-ing in ways that, I hope, even an externalist could accept. I try to show how one could take BS-ing to grant a person (external) justification for some belief without also conferring rational confidence in that belief.

William Alston was the first and most prominent externalist to argue that one can attain justification for a belief using BS-ing reasoning. He writes,

On my view, a belief is justified if and only if it is based on an adequate ground; that is, it is necessary only that the ground be adequate, not that the subject know or justifiably believe this...But then I can be justified in accepting the outputs of a certain doxastic practice without being justified in believing that the practice is reliable. [...] The argument would still be *epistemically circular*, for I am still assuming *in practice* the reliability of [the doxastic practice] in forming [the beliefs that are its outputs]. [...] [But] the epistemic circularity does not prevent justification from being transmitted from the premises to a conclusion that would have been unjustified except for this argument.<sup>24</sup>

Alston argues that all one needs to be justified in accepting the conclusion of a BS-ing argument is that one's BS-ing premise be based on an "adequate ground." One's ground for a belief is adequate if the doxastic process of which it is an output is *in fact* reliable. That is, one need not already know or justifiably believe that the process is reliable. The fact that the process *is* reliable gives one justification for believing the process's outputs, which one can

---

<sup>24</sup> William Alston, *The Reliability of Sense Perception* (Ithaca: Cornell University Press), 1993, pp. 16-7.

then use (along with the other needed premises) to gain justified belief in the conclusion that the process is reliable.<sup>25</sup>

Alston focuses primarily on the reliability of our perceptual experience in giving us justified beliefs about the physical world. Let's apply his account to the metaethical concerns that have been our focus. Recall

**PHENOMENAL:**

- (7) My moral experience presents  $P_1$ - $P_n$  as true or false in a mind-independent way. [objectivity interpretation of moral experience]
- (8)  $P_1$ - $P_n$  are true or false in a mind-independent way. [commitment to MI]
- (9) Thus, my moral experience is consistently accurate in indicating that MI is true.
- (10) Thus, my moral experience is a reliable source of beliefs about whether MI is true.
- (11) If a belief source is reliable, then one is justified in trusting it, i.e., in believing its outputs.
- (12) Thus, I am justified in believing MI on the basis of my moral experience.

Alston's version of externalism about epistemic justification is *reliabilism*, of which (11) is a roughly accurate characterization.<sup>26</sup> If we assume (11), you would be justified in believing (10) on the basis of (7) and (8), as long as (10) were in fact true. In other words, if we assume reliabilism, then as long as moral experience is a reliable source of metaethical beliefs, you would be justified in believing that  $P_1$ - $P_n$  are true or false in a mind-independent way on the basis of moral experience. From this, you could conclude that you are justified in believing that moral experience is metaethically reliable, and thus that you are justified in believing MI on the basis of your moral experience. You wouldn't need to know or justifiably believe that (10) is true *before* using PHENOMENAL. In using PHENOMENAL, you would be BS-ing by

---

<sup>25</sup> Alston seems to think the fact that the doxastic practice is actually reliable would not, on its own (i.e., without one's using the epistemically circular argument), justify one in believing that it is reliable. One would need to believe that the process is reliable *in virtue of* its generating the beliefs that it does.

<sup>26</sup> More precisely, a standard form of reliabilism holds that if  $S$ 's believing that  $p$  at  $t$  results from a causally reliable belief-forming process, then  $S$ 's belief in  $p$  at  $t$  is justified. See Alvin I. Goldman, "What is Justified Belief?" in G.S. Pappas (ed.) *Justification and Knowledge* (Dordrecht: Reidel, 1979), pp. 1-25.

presuming the truth of (10) in helping yourself to (8), but Alston argues that the fact that you have to BS to conclude (10) does not prevent you from being justified in believing it.

Can we find solace in this reliabilist alternative? I think not. Let's consider a version of the best-case scenario in which both (10) and (11) are true, but one knows (11)—i.e., one knows that reliabilism is the correct theory of justification—and does not (yet) know (10). As stated in §2, our main concern is to figure out whether to endorse MI on the basis of moral experience. To simplify the scenario, imagine that there are just three attitudes one could take toward the claim that moral experience is metaethically reliable: one could be skeptical, neutral, or confident. The reliabilist tells us, in effect, that if moral experience is a reliable source of beliefs about MI, then we are justified in believing MI on the basis of moral experience. Now, if a person is skeptical or neutral, reliabilism will not help her decide whether to believe MI. Since the skeptic and the neutral party are looking for some reason to believe the antecedent of the above conditional—i.e., that moral experience is a reliable source of beliefs about MI—they will not assume the consequent in the way they would need to—by BS-ing—in order for it to seem reasonable to accept the antecedent. Thus, they could not be rationally confident in MI by BS-ing in this way, even if they were justified in believing MI on the basis of moral experience (because moral experience was, in fact, metaethically reliable).

An externalist might object that this characterization of the epistemic defect of BS-ing either presupposes internalism or else employs a dialectical view of what can be achieved by BS-ing, as opposed to a normative one. While the normative view focuses on which beliefs a person is justified in holding, the dialectical view focuses on whether a person who does not already accept a conclusion could be rationally convinced of it via a BS-ing argument, given what else the person believes. This objection picks out an important feature

of my approach, but I do not think this feature is objectionable. For the purposes of this discussion, I have granted both reliabilism about justification and the metaethical reliability of moral experience. Then, I have asked whether accepting reliabilism would help the skeptic or the neutral party decide whether to believe MI on the basis of moral experience, and thus whether to attempt a defense of MI. I have argued that reliabilism does not help them, since BS-ing precludes them from having rational confidence in the relevant conclusion.

My approach appeals to dialectical considerations because these are essential to the process of reasoning about what to believe. They concern which kinds of argument can be used to move rationally from one set of beliefs to another, and they help to capture the perspective of the person who must decide what to think. Alston himself recognizes the challenge:

The [BS-ing] argument will not do its job unless we *are* justified in accepting its premises; and that is the case only if sense perception is in fact reliable. This is to offer a stone instead of bread. We can say the same of any belief-forming practice whatever, no matter how disreputable. We can just as well say of crystal ball gazing that if it *is* reliable, we can use a track-record argument to show that it is reliable. But when we ask whether one or another source of belief is reliable, we are interested in *discriminating* those that can be reasonably trusted from those that cannot...<sup>27</sup>

The inability of BS-ing reasoning to help one discriminate between reliable and unreliable belief sources is what prevents one from acquiring rational confidence in a conclusion by BS-ing. Put differently, since we cannot be rationally confident that we are justified in accepting the premises of a BS-ing argument, we cannot take heart that the BS-ing argument “does its job” of justifying us in accepting the conclusion.

Let’s now consider whether it makes a difference if the person is antecedently confident in the conclusion of the BS-ing argument. In other words, does one’s confidence, when leveraged through BS-ing, become *rational* confidence? It is difficult to see how it

---

<sup>27</sup> Alston (1993), p. 17.

could. When a person's antecedent confidence that moral experience is metaethically reliable leads her to BS, she accepts that  $P_1$ - $P_n$  are true or false in a mind-independent way due to that confidence, rather than due to the independent plausibility of this latter claim. Using the BS-ing argument does not seem to be helping her do any work—e.g., by showing that she should believe something she did not already believe, or by lending additional support to something she already believes—so it does not seem that her confidence in the conclusion that moral experience is metaethically reliable becomes any *more* rational than it was before. This empty result would seem to hold *even if* she were justified in believing that  $P_1$ - $P_n$  are true or false in a mind-independent way because, in fact, moral experience *is* metaethically reliable.

Committed externalists might not find the arguments of this section convincing. They are externalists for a reason, after all. Perhaps they would push back by denying that rational confidence is something we should care about. Or, if they accepted it as something we should care about, perhaps they would offer an interpretation of rational confidence that folded it into (externalist) justification. Against the first approach, I have argued that rational confidence is a kind of epistemic value we must care about, as agents deciding what to believe. Against the second approach, I have tried to show that rational confidence cannot be accounted for by a reliabilist conception of justification. If these two claims are correct, then embracing reliabilism will not allow us to avoid the central worries about what BS-ing can accomplish in defense of MI.

#### **4.2 Respectable Companions-in-Guilt?**

Nevertheless, defenders of externalism have a powerful rejoinder at their disposal. Recall that Huemer defends Phenomenal Conservatism on the grounds that a rejection of

the view would open the door to global skepticism. While Phenomenal Conservatism is a form of internalist foundationalism, externalists can marshal a very similar defense. Consider this passage from Selim Berker on the general challenge of trying to vindicate any basic source of belief:

...[There is a] general epistemological problem of how we can show that our most fundamental cognitive faculties (perception, introspection, induction, deduction, intuition—what have you) are reliable without relying on those very faculties when attempting to show this... There seems to be something viciously circular about appealing to a given cognitive faculty when attempting to vindicate the epistemic standing of that very faculty. But, with our most basic cognitive faculties, what recourse do we have except to appeal to those faculties during their vindication?<sup>28</sup>

The externalist could use the generality of this problem to offer the following **Epistemic Companions-in-Guilt** defense of MI (EPISTEMIC COMPANIONS): We cannot vindicate *any* of our basic capacities without BS-ing, so if we hold that a person lacks grounds for rational confidence in a conclusion reached (only) through BS-ing, we will have to accept a form of global skepticism according to which a person lacks grounds for rational confidence across the board. The externalist could then claim either that global skepticism is unacceptable on its face, or that articulating the skeptical view would be self-undermining. It would be self-undermining because we would have to rely on principles of inference, such as *modus ponens*, in order to state the argument, but we could only vindicate the use of those principles through BS-ing. Thus, it may not be possible even to articulate the view without falling into inconsistency. The externalist could claim, then, that because we should (or could) not accept global skepticism, we should (or could) not accept the view that reaching a conclusion via BS-ing prevents us from having rational confidence in it. In short, EPISTEMIC COMPANIONS argues that because BS-ing in favor of MI is on the same epistemic footing as

---

<sup>28</sup> Selim Berker, “Does Evolutionary Psychology Show that Normativity is Mind-Dependent?” in Justin D’Arms and Daniel Jacobson (eds.), *Moral Psychology and Human Agency: Essays on the New Science of Ethics* (Oxford, UK: Oxford University Press), p. 39. (My page citation here references the manuscript rather than the published version.)

BS-ing in favor of other basic sources of belief, we should not take the former to be especially problematic.

There may also be a practical analogue to this defense. The **Practical Companions-in-Guilt** defense of MI (PRACTICAL COMPANIONS) begins by noting that we could not, in practice, abandon relying on our basic sources of belief just because we lack a non-BS-ing vindication of those sources. Part of what makes any given belief source appropriately *basic* seems to be that we must trust or rely on it if we are to be rational agents at all. For instance, it is almost inconceivable to abandon introspection as a guide to our own occurrent mental states, or to abandon *modus ponens* as a principle of reasoning. For this reason, basic belief sources seem to retain a kind of immunity from practical disrespect. Thus, PRACTICAL COMPANIONS promises a form of protection for one's practical reliance on (the objectivity interpretation of) moral experience for believing MI. By lumping moral experience together with other basic sources of belief, we prevent it from being singled out for bullying.

Both EPISTEMIC COMPANIONS and PRACTICAL COMPANIONS claim that moral experience is a basic source of metaethical beliefs just like, say, perception is a basic source of beliefs about the physical world. But are these belief sources truly companions? Let's approach this question by considering how we treat skeptical challenges to these sources. Most philosophers seem to think that when it comes to skeptical challenges, we should normally default to common sense. And common sense is decidedly non-skeptical. One reason this approach might seem wise is that many skeptical challenges remain mere conceptual possibilities. While these challenges present obstacles to certainty, they don't have the power to undermine our faith that our beliefs are, in general, in pretty good shape.

This is largely because while we may lack non-BS-ing vindication of our basic belief sources, we also lack any positive reason to suspect that these skeptical possibilities are *actual*.

It may be possible, for instance, that I am massively deceived about the existence and character of the external world, since there is an evil demon playing tricks on me, or because I am a brain in a vat. But I don't seem to have any positive reason to suspect that I am so deceived. Likewise, I don't seem to have any positive reason to suspect that relying on *modus ponens* corrupts my reasoning, or that my mathematical intuition leads me to make systematic errors of arithmetic.

In the next section, I argue that we should not extend this policy of non-concern to moral experience. Unlike in the case of basic belief sources, we do have reason to suspect that moral experience is an unreliable guide to metaethics.

### 4.3 Undercutting Reliability

Let's return to our crop cultivation scenario. Imagine that you're on the way back to your hut, steeling yourself to sacrifice two more precious goats. Your neighbor notices your intent expression and asks what you're up to. You tell him that, on the shaman's advice, you are sacrificing your goats so that the gods will reward you with a successful harvest. Your neighbor looks puzzled. He says that the village has not yet installed the new shaman. He should know, since he is on the search committee. However, your neighbor has heard rumors of an unscrupulous knitter who has been thinning the local goat population in order to drive up the prices of her mohair sweaters. While the knitter has been suspected of killing the goats herself, your neighbor tells you, she may have found a new scheme to get others to do the dirty work for her. You are dismayed. You decide to hike up to the shaman's cave to find out for yourself. An hour later, your legs still burning from the steep ascent, you enter the cave's dank blackness. As your eyes slowly adjust, sure enough you spot her there, way in the back, lit by a single candle. She is knitting and laughing.

After you retreat from the cave, you decide to reconsider your earlier reliance on the knitter's advice. Suppose that you recognize in retrospect that your excitement at the thought that you had finally found a solution to your agricultural troubles led you to BS your way to the conclusion that the knitter is reliable. Also suppose that you agree with my claim that one cannot be rationally confident in a conclusion reached (only) through BS-ing. Finally, suppose that you're attracted to an externalist theory of justification. You reason that it still might be *true* that the knitter is a reliable source of beliefs about crop cultivation. And if it were true, you could use CROP to gain justification for the belief that the knitter is a reliable source of such beliefs. Imagine that, unbeknownst to you, your crops have been failing because your goats roam out of their pen at night and eat your seedlings. In this case, the knitter would be accidentally reliable—sacrificing your goats *would* lead to a more successful harvest—even though you would have no grounds for rational confidence in the claim that she is reliable. If externalism were true, your lack of grounds for rational confidence in this case would not deprive you of justification. So perhaps you should maintain your belief in the knitter's reliability. You might try to bolster the case for sticking with this belief by reflecting on the fact that, as EPISTEMIC COMPANIONS and PRACTICAL COMPANIONS argue, we cannot abandon BS-ing vindications wholesale.

After running through these considerations, should you maintain your trust the knitter? Of course you should not. For now you have an undercutting defeater for the claim that the knitter is a reliable source of beliefs about crop cultivation. Whereas before you assumed that the knitter had supernatural and agricultural insights that she wanted to share for your benefit, now you realize that you have no reason to believe that she has such insights, and that she has only been trying to benefit herself by convincing you to do something that (she believes) will leave you worse off. Thus, the problem now is not simply

that you lack grounds for rational confidence in the knitter's reliability, but that you also have solid grounds *to doubt* her reliability. It seems clear, then, that you should withdraw your trust in the knitter and save your goats.

Let's call the epistemic principle guiding this last judgment

**Dubious Reliability:** If one has solid grounds to doubt the reliability of a belief source and only a BS-ing argument in favor of its reliability, one should not rely on it.<sup>29</sup>

Dubious Reliability aims to capture the intuitive idea that some instances of BS-ing are more suspect than others. All instances of BS-ing face the main problem we have been discussing for much of this chapter: when a person has only a BS-ing argument in favor of some conclusion, using the argument does not generate grounds for rational confidence in the conclusion. As we have seen, however, some philosophers claim that we can take comfort in the recognition that all of our basic belief sources are equally dogged by this problem, such that it would be either impossible or unreasonable to take ourselves to have epistemic reason to cease relying on those sources. Depending on one's broader epistemological sympathies, one could articulate this idea in terms of having (externalist) justification for relying on those sources, or in terms of having an overarching reason not to abandon confidence in those sources.

---

<sup>29</sup> A person's grounds are solid if they are based on good evidence and do not neglect other relevant evidence. One worry: someone might think that this definition of solid grounds entails that the second conjunct of Dubious Reliability—"...and only a BS-ing argument in favor of its reliability"—does no work for the principle, since a defender of my view would treat a BS-ing argument as no grounds for rational confidence that some belief source is reliable. By the same token, someone who defends the use of BS-ing arguments to gain justification would deny that the antecedent of Dubious Reliability ever obtains, since the truth of its second conjunct would always entail the falsity of its first conjunct. In other words, if one has a good BS-ing argument, then, by definition, one's grounds to doubt the reliability of the belief source would not count as *solid*. However, this way of responding to Dubious Reliability faces the same problem we have been discussing—the externalist gives us no way of distinguishing a “good” BS-ing argument from a bad one. In the absence of such a feature, we must figure out some practical way to determine which BS-ing arguments we should pardon in light of companions-in-guilt style defenses, and which we should reject due to their independent implausibility. Dubious Reliability aims to help us with this task.

Nevertheless, some instances of BS-ing cannot mount this companions-in-guilt style of defense, since they seek to vindicate belief sources that we also have solid grounds to doubt are reliable. It is not just that we lack a non-BS-ing case for these sources; we also have a good case against them. Dubious Reliability claims that we should abandon reliance on such sources of belief. If Dubious Reliability is sound, then your discovery of the knitter's nefarious plot should lead you to stop relying on her advice. I will now argue that an analogous pair of considerations should lead one not to rely on (the objectivity interpretation of) moral experience to believe MI.<sup>30</sup> I will discuss two reasons why people would be prone to experience moral demands as objective even if they were not objective. If these explanations are satisfactory, they serve as undercutting defeaters for the claim that moral experience is metaethically reliable.

The first proposed undercutter appeals to an evolutionary explanation of why it would be adaptive for people to experience moral demands as objective.<sup>31</sup> The overarching idea is that we are disposed to experience morality as objective largely because having this disposition, or some proto-moral version of it, enhanced the reproductive fitness of our ancestors. But unlike with other capacities and dispositions shaped by evolutionary forces—such as sense perception or mathematical reasoning—we have no evolutionary rationale for thinking that the dispositions influencing our moral experience are reliable guides to

---

<sup>30</sup> Recall that PHENOMENAL begins with the premise that moral experience presents moral beliefs in which one is confident as true or false in a mind-independent way. This is what I have called the objectivity interpretation of moral experience, and it is what enables the proponent of PHENOMENAL to BS to the conclusions, first, that moral experience is metaethically reliable, and, second, that one is justified in believing MI on the basis of moral experience.

<sup>31</sup> Caveat: I can at most sketch the outline of such an explanation here. A full account would need to address a number of complicated issues regarding exaptation, the levels of selection, the relationship between exhibiting pro-social behaviors and acting on genuine moral beliefs, among much else. Nevertheless, I hope this sketch is sufficiently plausible to make my point.

metaethical truth.<sup>32</sup> If we combine this claim with the charge of BS-ing, we can appeal to Dubious Reliability to argue that we should not rely on (the objectivity interpretation of) moral experience to believe MI. Both the nefarious knitter and evolutionary forces shape some of our beliefs in unreliable ways. Their purposes—the knitter’s self-interest in the first case, our ancestors’ reproductive fitness in the second—are unconnected to the truth of the beliefs they aim to produce in us. So, absent a plausible reason to think there is a fortunate alignment between those beliefs and what is true, we should not rely on these sources.<sup>33</sup>

Here is a general template for making the case that natural selection would favor the disposition to interpret moral norms as objective, even if they were not. Acting in accordance with accepted moral norms, or their proto-moral progenitors, was fitness enhancing because it allowed our ancestors to participate in relationships of mutually beneficial cooperation. However, given the appeal of self-interest, our ancestors were more likely to act in accordance with moral norms if they viewed their normative force as independent of and overriding with respect to their self-interested motives. And a very natural way to interpret the notions of independence and overridingness is in terms of objectivity. Objective moral norms would apply to us regardless of whatever other desires, plans, or inclinations we might happen to have.

---

<sup>32</sup> I don’t want to suggest that our rationales for trusting sense perception or mathematical reasoning are purely evolutionary. Rather, the point is that the evolutionary accounts of these capacities’ development make it implausible that we would be disposed to believe their outputs even if they were false. (This claim would need to be rather restricted in the case of sense perception, which allows us to represent the world in ways that are useful but greatly simplified.) I am claiming that the same is not true when it comes to treating moral experience as a source of metaethical beliefs—an evolutionary account explains why we would be disposed to believe MI even if MI were false.

<sup>33</sup> The analogy is imperfect in two ways. First, evolution has no aims or purposes, though we can reconstruct an explanation for why certain traits were adaptive in certain contexts, and so why they were selected for. Talk of “purposes” or “aims” is thus merely suggestive. Second, I refer to “evolutionary forces” as the unreliable source in order to draw an evocative parallel with the nefarious knitter. It would be more accurate, though also wordier, to refer instead to “moral experience as shaped by evolutionary forces,” since (the objectivity interpretation of) moral experience is the source with which we are concerned.

We can fill in the details of this template in a variety of ways. For instance, Brian Skyrms has worked at the intersection of rational choice theory and evolutionary dynamics to show how dispositions to cooperate can become common in a population.<sup>34</sup> For cooperation to constitute an evolutionarily stable strategy in a population containing other cooperators, one must learn to identify defectors and to develop dispositions to blame or otherwise punish them when they transgress the norms that support the cooperative scheme.<sup>35</sup> Temptations to defect would be strong, since doing so would usually earn the defector a higher payoff in fitness if the other party continues to cooperate. On the other hand, the costs of being caught defecting could be substantial. A defector might face violent retribution or the prospect of being outcast from the group. Even developing a bad reputation could be devastating, since it would make others less willing to cooperate with one in the future. Therefore, the members of such a population will be best served by developing mechanisms to identify and blame defectors, as well as to avoid being identified as defectors themselves.

In order to secure the benefits of cooperation, then, one needs a way of convincing others that one will not be swayed by the temptation to defect. Robert Frank suggests that the development of moral sentiments satisfies this need, since their display can serve as a signal that one is trustworthy. He writes,

Moral sentiments may be viewed as a crude attempt to fine-tune the reward mechanism, to make it more sensitive to distant rewards and penalties in selected instances... People with genuine moral sentiments are better able than others to act in their own interest.<sup>36</sup>

---

<sup>34</sup> Brian Skyrms, *Evolution of the Social Contract* (New York: Cambridge University Press, 1996).

<sup>35</sup> A defector is a free-rider within a cooperative scheme who attempts to conceal his free-riding.

<sup>36</sup> Robert H. Frank, *Passions within Reason: The Strategic Role of the Emotions* (New York: Norton, 1988), pp. 90, 91.

Frank's claim is that humans developed "genuine moral sentiments" as a roundabout way of enabling them to act prudentially. This claim would be more plausible if it referred to the expected payoff in terms of reproductive fitness, rather than in terms of a person's self-interest in the normal sense. For Richard Dawkins may be right that selection occurs at the level of the gene, rather than at the level of the individual organism or group. If that claim is correct, then some trait could be adaptive by promoting genetic fitness while being detrimental to an individual's interests.<sup>37</sup>

But the heart of Frank's claim seems plausible: it is by genuinely caring about adhering to moral standards that one is able to reap the benefits of participating in cooperative schemes, because the most reliable way to seem like one cares is *to actually care*. The development of a conscience enables one's concern for adhering to moral standards to guide one's actions. The feeling of being bound to act morally and the prospect of experiencing guilt if one falls short of doing so can motivate a person to act in ways that would otherwise seem to contravene the dictates of self-interest. Just as importantly, it can signal to others that one is unlikely to defect even when others are not well placed to detect defection, thereby engendering the trust necessary for cooperation. In order for one's conscience to perform this function, the moral standards to which it appeals cannot seem like mere desires or preferences. Rather, one's conscience must present moral standards as *overriding* one's mere desires and preferences. As I have claimed above, it is quite natural to conceive of this overridingness and independence in terms of objectivity.

If this evolutionary explanation is successful, it serves as an undercutting defeater for the claim that moral experience is metaethically reliable. It shows how (the objectivity interpretation of) moral experience has evolved to enhance reproductive fitness, rather than

---

<sup>37</sup> Richard Dawkins, *The Selfish Gene* (New York: Oxford University Press, 1976).

to represent the metaethical truth. Of course, MI might nevertheless be true, and moral experience might nevertheless be metaethically reliable. But you cannot responsibly rely on it. For in addition to having only a BS-ing argument for the metaethical reliability of moral experience, you also have good reason to doubt its reliability, since you would be disposed to believe MI even if MI were false.

The second proposed undercutter claims that moral experience is not metaethically reliable because it is easy to misinterpret. In particular, it claims that the objectivity interpretation of moral experience is mistaken, since the experienced independence of moral norms is better explained in terms of their *demandingness*. When we accept certain types of moral norms, we commit to giving those norms a form of *deliberative priority* over other considerations that figure in our decisions about what to do, such as our preferences, desires, and plans. We thereby experience those norms as independent of these other subjective considerations. For instance, we think the wrongness of pillaging takes priority over the enjoyment we might gain from running around and setting things on fire.

The natural next question asks *why* we grant moral norms this deliberative priority. In Chapter 3, I will defend an answer that ties our acceptance of these norms to the practical authority we grant to others within relationships of mutual recognition. The normative force of moral claims is thus independent of a given subject in a second way: it comes from the practical authority of other people with whom one interacts, and is expressed by their standing to make moral demands. In sum, I will argue that the experienced independence of moral norms can be explained by appealing to the practical authority of others, our practical commitment to granting them that authority by participating in relationships of mutual recognition with them, and the demandingness of the norms that structure these relationships.

If these two proposed undercutters defeat the metaethical reliability of moral experience, then three of the considered approaches to defending MI fail. First, EPISTEMIC COMPANIONS and PRACTICAL COMPANIONS both fail, since we have shown that moral experience is not on a par with our other, less suspect basic belief sources. We have positive reasons to doubt the metaethical reliability moral experience, but not positive reasons to doubt, say, the mathematical reliability of our mathematical faculties. Second, the approach of turning to externalism fails, given that we have reasons to doubt the metaethical reliability of moral experience. Third, the approach of appealing to Phenomenal Conservatism fails, since that view defends reliance on appearances only in the absence of defeaters.

## 5. Conclusion

In this chapter, I have advocated a skeptical challenge for defenders of MI. We can undermine our grounds for rational confidence in MI by reflecting on the unreliability of the process that led us to find the view plausible, and so worth defending, in the first place. I have criticized the claim that the way things seem when we are engaged in moral thinking gives us good reason to conclude that morality is objective. I have tried to show how operating from that rationale requires one to BS, and I have argued that BS-ing fails to generate grounds for rational confidence in a belief source when one also has solid grounds to believe the source is unreliable. I have appealed to two such grounds. First, I have claimed that evolutionary considerations explain why we would be disposed to interpret morality as objective even if it were not objective. Second, I have claimed that moral experience is easy to misinterpret, and briefly sketched my favored alternative interpretation: the demandingness of moral concepts reflects their function in interpersonal relationships to which we are practically committed.

I have argued, in short, that we should cut the BS in metaethics. If my arguments have been successful, then those who would defend MI will need to appeal to a more compelling rationale. In Chapter 2, I turn to one promising candidate to fill that role: the view that claims about the objectivity of morality are nothing more than abstract, first-order moral claims, and so can be assessed only on first-order moral grounds.

## ON TREATING MORAL OBJECTIVITY AS A MORAL VIEW

### 1. Introduction

In this chapter, I discuss a view that rejects the general approach I have so far pursued to evaluating belief in MI. Let's begin by reflecting on the assumptions that underlie the skeptical challenge advocated in Chapter 1, so that we may better understand the differences between the two approaches.

The skeptical challenge claims that we can undermine rational confidence in MI by reflecting on the unreliability of the process that led us to find that view appealing, and so worth defending, in the first place. This kind of challenge is familiar from other contexts. If I discover that the only reason I believe I'm destined to become a famous singer is that I've had the belief implanted through hypnosis, or because a fortune-teller told me I'd one day perform in front of fainting teens and uncomfortable parents, that discovery undermines my rational confidence in the belief. This approach to motivating a skeptical result appeals to descriptive considerations about how beliefs, or dispositions to believe, have been formed in order to cast doubt on the rationality of people's holding those beliefs.

But the famous singer case is simpler than the MI case, since the former concerns only beliefs about descriptive matters—what is or will be true in the natural world—and appeals to relatively uncontroversial epistemic principles in assessing those beliefs. By contrast, when we assess belief in MI we must include considerations from both descriptive and normative domains. And, at least implicitly, we must take positions on how these domains relate to one another, as well as on which epistemic principles govern our views about those relations. These issues are considerably more complicated.

The approach I have advocated suggests that in trying to understand the nature of moral normativity, we face what I'll call the

**Reconciliation Problem:**

How do we reconcile the perspective we inhabit when engaging in moral thought with the perspective we inhabit when studying ourselves—including our dispositions to engage in moral thought—as parts of the natural world?

Reflecting on our experiences from these two perspectives can lead us to form two different conceptions of our moral lives. The problem is that the first, *practical conception* proves difficult to reconcile with the second, *naturalistic conception*, yet we can abandon neither.

The practical conception of our moral lives aims to capture what we are doing when we make moral claims, blame someone for falling short of what she ought to do, or feel guilty about falling short ourselves. This conception attempts to make sense of our experience of confronting considerations that bear on how we should act, such as by giving us moral reasons not to  $x$ , or by obligating us to  $y$ . I explained in Chapter 1 how a natural way of interpreting moral experience could lead one to adopt a conception that treats morality as objective. I argued that we lack grounds for rational confidence in that conception, in part because a naturalistic conception of our moral lives explains why we would be disposed to conceive (practically) of morality as objective even if it were not. Thus, we can pose the reconciliation problem in a more pointed way: our naturalistic conception of our moral lives can interfere with our ability to fully endorse our practical conception of our moral lives. Chapter 1 argues, in effect, that several prominent rationales in favor of MI leave us unable solve the reconciliation problem.

In appealing to empirical considerations to undermine rational confidence in MI, I have also presupposed a second view. Let's call it

**Empirical Relevance:**

Empirical considerations can be relevant to how we understand the nature and status of morality.

The skeptical challenge relies on Empirical Relevance to argue that several widespread approaches to MI fail to solve the reconciliation problem.

To be clear, the skeptical challenge does *not* claim that our practical perspective is inherently incompatible with our naturalistic perspective. Accepting that claim would be tantamount to accepting nihilism.<sup>38</sup> Nevertheless, nihilism is the specter that the reconciliation problem presents us with and that constructive metaethical theories ward against. The skeptical challenge appeals to empirical considerations to argue for a weaker conclusion: we lack grounds to be rationally confident in *one particular conception* of what we're doing when engaging in moral thought. If we assume that our naturalistic understanding of ourselves is not open to radical reinterpretation, the skeptical challenge presents us with three main options. We can abandon hope of reconciliation, we can attempt to revise our practical conception, or we can attempt to explain why our two conceptions, as originally conceived, are not actually in tension.

Let's briefly consider what each approach would involve. The first option is to argue for nihilism. A proponent of this option would try to show that the most plausible conception of what moral thought involves makes the reconciliation problem unsolvable. The success of my case for Intersubjectivism—or, for that matter, the success of any other constructive metaethical theory—would show that this option is misguided, so I set it aside for now. The second option is to argue for a form of MD that makes our practical

---

<sup>38</sup> This statement assumes that if a person determined that the views from our practical and naturalistic perspectives were somehow incompatible, she would conclude that our practical perspective is thereby discredited. I am aware of nearly no one who holds the opposite view, i.e., that the conflict would discredit our naturalistic perspective, with Thomas Nagel being a rare exception (see his *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False* (New York: Oxford University Press, 2012)). Structurally similar views have been advocated in other contexts. For example, some have held that a conflict between scientific evidence and religious doctrine—say, about the age of the Earth—gives us decisive reason to distrust the former.

conception easier to reconcile with our naturalistic conception. The main challenge for this approach is to show that the revised practical conception can make sense of moral thought's essential features. I attempt to meet this challenge in Chapter 3. The third option is to defend MI, and one could undertake this project in two distinct ways. One could offer a rationale for MI that solves the reconciliation problem, or one could offer a rationale that dissolves the problem by explaining why, despite appearances to the contrary, it does not exist.

The view I discuss in this chapter explores the last possibility. It denies that we face the reconciliation problem because it rejects Empirical Relevance. More precisely, the view endorses

**Substantive Stance:**

Any claim that is relevant to how we understand the nature and status of morality is itself a substantive moral stance.

Substantive Stance entails that most of what has traditionally passed for metaethical debate is really just moral debate, and so is radically misunderstood by the majority of those who engage in it. The metaphysical, epistemological, and psycho-linguistic worries about MI, along with the alternative views those worries motivate, are not claims "external" to moral discourse capable of undermining its purported objectivity from some theoretical remove. Rather, they are general *moral* claims about what morality is or would have to be like, and are thus "internal" to moral discourse. As such, these claims must be assessed on moral grounds.

One can appeal to Substantive Stance in defending either MI or MD. However, the view fits much more naturally with a defense of MI, for it offers the prospect of treating morality as independent in two, complementary ways. The truth of Substantive Stance would enable one to treat moral facts as ultimately independent of people's attitudes and practices

(MI), and to treat this very claim about moral facts as independent of any non-moral factors that might mistakenly be thought capable of vindicating or undermining it (Substantive Stance). In this way, joining MI with Substantive Stance allows one to make an expansive claim about moral objectivity that, at the same time, insulates itself from skeptical attack. If this move were successful, it would dissolve the reconciliation problem by showing it to be illusory. There would be no genuine conflict between our naturalistic and practical perspectives because only the latter would be relevant to assessing morality's nature or status. Accepting Substantive Stance would allow us to treat the skeptical challenge as a form of substantive moral skepticism, which we could reject either because it was self-contradictory or because it proved less plausible than other substantive moral claims with which it conflicted.<sup>39</sup>

My goals in this chapter are to consider the arguments for Substantive Stance put forward by Ronald Dworkin and to assess how those arguments impact the skeptical challenge.<sup>40</sup> I discuss Dworkin's general view in §2 and his arguments for Substantive Stance in §§3-4. I conclude in §5 by considering the upshots for Intersubjectivism.

## **2. Dworkin on the Varieties of Skepticism**

### **2.1 Dworkin's View and Apparent Alternatives**

Ronald Dworkin appeals to Substantive Stance to defend MI. In so doing, he attempts to dissolve the reconciliation problem by rejecting Empirical Relevance. Dworkin argues that the moral domain is *independent*, which entails that any claim seeking on its own to vindicate or undermine morality must itself be a moral claim. Thus, he claims that attempts

---

<sup>39</sup> I say more about these two possibilities below.

<sup>40</sup> Ronald Dworkin, "Truth and Objectivity: You'd Better Believe it," *Philosophy and Public Affairs*, Vol. 25, No. 2 (Spring 1996), pp. 87-139. Dworkin, *Justice for Hedgehogs* (Cambridge, MA: The Belknap Press of Harvard University Press), 2011.

to raise skeptical doubts about MI on non-moral grounds—of the kind I pursue in Chapter 1—are a sham. But, by the same token, so are attempts to *defend* MI from such doubts by responding on non-moral grounds, as do most advocates of MI. Dworkin thus courts enemies on both sides of the aisle, with obvious delight.<sup>41</sup>

Dworkin begins by articulating what he calls the *ordinary view*. This is the view of morality that he believes most non-philosophically-inclined people hold. And, coincidentally, he believes it is the correct view. The ordinary view has three main tenets. First, it maintains that there are objective truths about what is right or wrong, truths that do “not depend on what anyone thinks or feels.”<sup>42</sup> Thus, the ordinary view advocates MI. Second, it advocates the epistemic view that moral argument—not any kind of moral revelation, perception, or intuition—is the only way to discover these objective moral truths. Third, the ordinary view holds that

...General questions about the basis of morality—about what makes a particular moral judgment true—are themselves moral questions. Is God the author of all morality? Can something be wrong even if everyone thinks it is right? Is morality relative to time and place? [...] These are abstract and theoretical questions, but they are still moral questions. They must be answered out of moral conscience and conviction, just like more ordinary questions about right and wrong.<sup>43</sup>

This third tenet is an endorsement of Substantive Stance.

Not much hangs on this terminological issue, but it would be more accurate just to call this *Dworkin's view* rather than the *ordinary view*, since it seems unlikely that any “ordinary” person has ever held it. Consider the first tenet. Many ordinary people are unsure about whether their moral convictions *are* objectively true, though they usually go on acting from those convictions just the same. Perhaps these people would count as “philosophically-

---

<sup>41</sup> Here is one representative passage: “We cannot escape from morality’s independence, no matter how strenuously we struggle. Every effort we make to find a trap door out of morality confirms that we do not yet understand what morality is.” Dworkin (2011), p. 39.

<sup>42</sup> Ibid., p. 27.

<sup>43</sup> Ibid., p. 28.

inclined,” though, and so would no longer count as “ordinary” in Dworkin’s sense.

Regarding the second, epistemological tenet, many ordinary people believe that moral truths are revealed through people’s interactions with supernatural beings, or through scriptures based on those interactions, and that willingness to engage in moral argument betrays a morally problematic lack of faith. Many other people probably hold no beliefs about what makes moral claims true, and could not offer much in the way of argument for their moral convictions. Others probably believe that they *can* perceive moral truth, but, now taking up the third tenet, might not think that this epistemological belief is itself a moral conviction, or that there is any moral upshot if that belief is mistaken. The main point here is two-fold: what Dworkin calls the ordinary view is just Dworkin’s view, and any support he hoped to rally for that view by aligning it with common sense is illusory.

Dworkin calls those people *skeptics* who reject the first tenet of the ordinary view, the idea that moral convictions can be objectively true. Dworkin distinguishes several types of skepticism. First, he distinguishes *internal* from *external* skepticism:

Internal skepticism about morality is a first-order, substantive moral judgment. It appeals to more abstract judgments about morality in order to deny that certain more concrete or applied judgments are true. External skepticism, on the contrary, purports to rely entirely on second-order, external statements about morality [...and so...] is supposedly Archimedean: it stands above morality and judges it from the outside...[These skeptics] are able to denigrate moral truth, they say, without relying on it.<sup>44</sup>

One example of internal skepticism would be a view that denies that a wife owes special duties of obedience to her husband (the more concrete judgment), on the grounds that requiring such obedience would be inconsistent with respecting the wife’s autonomy (the more abstract judgment). This form of skepticism is internal because it rejects some moral claim(s) on the basis of accepting some other moral claim(s).

---

<sup>44</sup> Ibid., pp. 31-2.

There are two types of external skepticism. *Error skepticism* holds that all moral judgments are false because there are no “moral entities,” or truth-makers in virtue of which moral judgments could be true. As the name suggests, Dworkin has in mind views such as John Mackie’s error theory. Mackie argues that moral values would have to be objectively prescriptive to exist, but since there is nothing objectively prescriptive, there are no moral values, and so we are in error to the extent that our practices presuppose that there are.<sup>45</sup>

The second type of external skepticism is *status skepticism*, which holds that moral judgments play some function other than to describe how things are by picking out moral facts. Dworkin has in mind such views as expressivism, emotivism, prescriptivism, and quasi-realism, which characterize the making of moral judgments in terms of some non-(wholly-)cognitive states rather than in terms of beliefs about what is morally true. He notes that more recent versions of status skepticism have attracted proponents due to their promise of allowing people to be skeptical of the *status* of their moral convictions, while nevertheless going on to treat the *content* of those convictions just as seriously as do the virtuously non-skeptical ordinary people.

## 2.2 Three Worries

Before proceeding, I’d like to raise three worries about Dworkin’s treatment of moral skepticism. First, he claims that all forms of moral skepticism deny that any moral claims are true, or even truth-apt. This is a significant mistake. The skeptical challenge from Chapter 1, for instance, claims that we lack grounds for rational confidence in MI, but makes claims neither about whether MI is true nor about whether substantive moral claims are true

---

<sup>45</sup> To nitpick a bit: an error theorist would likely count some moral judgments true, e.g., “It is not the case that there is moral reason either to lie or not to lie.” Thus, Dworkin would be better off revising his characterization of error skepticism to hold that all judgments positing the existence of moral truths are false (or otherwise in error, depending upon how one wants to classify claims that suffer from presupposition failure).

or truth-apt. Other classic forms of skepticism take a similar approach: they challenge the epistemic status of some class of beliefs, given the quality of our reasons for holding those beliefs, but stop short of making the stronger claim that those beliefs are false or not apt for truth. By attributing to his skeptical opponents a stronger stance, Dworkin tips the rhetorical balance in his own favor by arguing only against the most extreme forms of skepticism. In the next section, I aim to show how this mistake undercuts the power of his arguments, especially with respect to my skeptical challenge.

The second worry concerns Dworkin's discussion of internal skepticism. His characterization of the view makes it clear that we are all internal skeptics about a potentially infinite number of moral claims—basically, any moral claim we would reject on the basis of taking a stance on some other, more general moral claim. For instance, I am an internal skeptic about the obligatoriness of spinning three times upon entering a room, and you are probably an internal skeptic about the blameworthiness of winking at puppies. By contrast, the term “skepticism” is usually reserved for views that express doubt about the status of a *whole class* of claims—regarding the existence of other minds, say, or the existence of the external world—so it is somewhat jarring that Dworkin uses it to describe the simple act of rejecting one moral claim on the basis of accepting another moral claim.

A more plausible model for internal skepticism is not these classic positions, but rather a more limited position like climate change skepticism. Most climate change skeptics do not take themselves to be skeptical about evidence-based inquiry as a whole, but rather purport to have special reasons for doubting the scientific consensus around climate change. For instance, they appeal to the unreliability of climate models given the natural fluctuations in temperature patterns over time, or they accuse scientists of allowing their political agendas to shape their interpretation of relevant data. In these ways, most climate change skeptics

make their claims, however misguided, from *within* the normal discourse of assessing the rationality of grounds for empirical beliefs. Since Dworkin's end game is to argue that all purportedly external claims are really just abstract internal claims, he has good dialectical reason to characterize internal skepticism using this kind of model: it clearly involves "internal" claims but also has a form that may be sufficiently general to accommodate the more familiar, purportedly external claims of moral skepticism.

The third worry is that Dworkin's characterization of MI's opponents is insufficiently comprehensive. He initially characterizes skepticism as the view that (some class of) moral convictions cannot be objectively true. He then claims that there are two forms of purported external skepticism: error skepticism, which holds that all moral claims are objectively false, and status skepticism, which holds that moral claims are not even apt for objective truth or falsity. But this characterization neglects a large family of views that treat moral claims as capable of being true (or otherwise correct) in some *non-objective* sense.<sup>46</sup> Such views include certain forms of subjectivism, relativism, and constitutivism, as well as Intersubjectivism.

While advocates of these views interpret them as alternatives to MI, Dworkin would interpret them as versions of *internal* skepticism. He would therefore interpret these views as making general moral claims about the conditions under which more particular moral claims are or can be true, as well as counterfactual moral claims about the conditions under which

---

<sup>46</sup> Dworkin is inconsistent here. For instance, he sometimes refers to external status skepticism as the view "that moral judgments do not even purport to be true" (Dworkin (2011), p. 40). There are three problems with this characterization. First, by failing to mention *objective* truth as the crucial notion, Dworkin fails to distinguish between views like expressivism, which treats moral judgments as non-truth-apt, from views like (certain forms of) relativism, which treats moral judgments as capable of being true or correct, but not objectively so. Second, and relatedly, this tacit equating of truth with objectivity may bias the reader against possible alternatives to the "ordinary view." Third, there may be important differences between what moral judgments purport to be and what they are. Dworkin means to focus on what they are.

other particular moral claims are false, e.g., whenever such claims presuppose objectivity. What would doom these views, according to Dworkin, is that the counterfactual claims they entail would be morally implausible. For instance, a simple form of subjectivism would entail a counterfactual such as “If I didn’t hold the conviction that murder is wrong, then murder would not be wrong.” Since this claim is morally misguided, we should reject the view that entails it.

One interesting wrinkle is that Dworkin uses this counterfactual-reading strategy to object to *any* view that deviates from his own. This fact suggests that what he thinks is ultimately objectionable about skeptical views is not their skepticism, but rather something they share with many other views, including most versions of MI: their denial of Substantive Stance. Any view about the nature of morality that can be reinterpreted so that it makes a claim of the form “If condition  $x$  were not met, then moral claim  $y$  would not be true” would be subject to Dworkin’s counterfactual-reading strategy. And the more abstract condition  $x$  is, the more likely the counterfactual claim would seem morally misguided.

For instance, consider a generic form of MI that posits ontologically substantial moral truth-makers. Dworkin argues that we would have to understand these truth-makers as something like moral protons—what he derisively calls “morons.” And he suggests that anyone who believes in such things deserves to be classified as a moron in the more usual sense of the term, though he neglects to mention that this group includes nearly every philosophical defender of MI going back to Plato. Dworkin would treat their views as entailing counterfactual moral claims such as “If there were no ontologically substantial moral truth-makers, then murder would not be wrong.” Since Dworkin thinks that the wrongness of murder depends solely on the moral case that can be made for its wrongness, he would interpret this counterfactual claim as seriously morally misguided. What makes

murder wrong is the way it harms people, treats people as mere means to one's ends, and so on, and therefore has nothing to do with fanciful metaphysical entities. In this respect, Dworkin thinks that other forms of MI misunderstand morality just as badly as do forms of MD.

Clearly, then, Substantive Stance is of central importance to Dworkin's view. In the next two sections, I take a closer look at Dworkin's arguments for Substantive Stance, with an eye to assessing how they impact the skeptical challenge from Chapter 1.

### **3. Dworkin against External Error Skepticism**

#### **3.1 The Main Arguments**

Dworkin's master argument against external skepticism has two main components. The first component appeals to Substantive Stance to argue that both forms of (purportedly) external skepticism are self-defeating because they are actually forms of internal skepticism. They are moral views that claim no moral view is objectively true. But, Dworkin insists, a view must claim objective truth for itself. Thus, external skepticism collapses in a contradictory mess. The second component is connected to the counterfactual reading strategy just discussed. This component relies on Substantive Stance to argue that even if external skeptical views somehow avoid the specter of self-defeat, once they are appropriately recast as forms of internal skepticism, we can see that they make or entail moral claims that should clearly be rejected in favor of the more plausible moral claims licensed by the ordinary view. As both components of Dworkin's master argument depend upon Substantive Stance, in this section I turn to considering the case for that view.

The first problem with Substantive Stance is that it appears just as vulnerable to self-defeat as Dworkin claims external skepticism is. After all, Substantive Stance denies that

there are any genuine metaethical claims, but seems *itself* to be a metaethical claim. It concerns the nature and status of certain purportedly moral claims, but does not make a substantive moral claim itself. After all, what substantive conclusions could one draw from the claim that any claim about the nature or status of morality is itself a substantive, moral stance? If the answer is “none,” then, absent some remediating clarification, Substantive Stance contradicts itself, and should be abandoned in favor of admitting that there are some genuine metaethical claims. In what follows, I will not rely on this objection, but rather treat it as one of several reasons to doubt that Dworkin’s case against external skepticism is as convincing as he takes it to be.

Dworkin’s primary approach to arguing for Substantive Stance involves testing our intuitions about whether and when there is a difference between internal and external claims about morality. He asserts, “Philosophy can neither impeach nor validate any value judgment while standing wholly outside that judgment’s domain. Internal skepticism is the only skeptical game in town.”<sup>47</sup> Dworkin first aims to convince us of this claim with regard to external error skepticism by offering an analogy in which a person denies there are such things as unicorns. Is this person making an external error claim *about* the discourse of unicorn zoology—that it is all bunk—or is she making a claim *within* that discourse?

Dworkin argues for the latter interpretation:

We may say that no claim anyone makes about the shape or color of unicorns is true because there are no unicorns. But we can’t then declare that no proposition of unicorn zoology can be true.<sup>48</sup>

The true proposition within unicorn zoology is presumably that there are no unicorns. But perhaps Dworkin also thinks the person who denies the existence of unicorns thereby commits herself to the claims there are no dappled unicorns, no unfriendly unicorns, no

---

<sup>47</sup> Ibid., p. 35.

<sup>48</sup> Ibid., p. 41.

slothful unicorns, and so on, ad infinitum. Instead of rejecting the discourse of unicorn zoology as she intended, then, the unicorn-denier has become yet another unwitting unicorn zoologist.

There are two problems with Dworkin's discussion here. First, it seems at least equally plausible that the unicorn-denier has made her claim "external" to the unicorn zoology discourse, perhaps as a biological claim about which animals we have evidence for—and therefore about which more particular zoological discourses it makes sense to posit or engage in—or else as a general physical or metaphysical claim about which things exist. If this interpretation is at least as plausible as Dworkin's, then the unicorn analogy provides no support for Substantive Stance. Second, it is difficult to grasp what, if anything, is at stake in classifying the rejection of unicorns as either an internal or an external claim. If this example were presented in isolation, it seems plausible that one's reaction would be to offer an initial intuition about the case, but then follow it up by asking "Anyhow, what does it really matter?" Perhaps there's some advantage to minimizing the number of distinct discourses we recognize, but it's not clear what that would count for. It is far from obvious, then, that we should use our intuitions about how to classify claims about unicorns as any kind of evidence for how we should classify purportedly metaethical claims, since the stakes are, by hypothesis, clearer and more significant for the latter classification.

Thankfully, Dworkin's next attempt appears less flimsy. He asks us to consider the following "conversation" as an illustration that there is no fundamental difference between internal and external moral claims:

A: Abortion is morally wicked: we always in all circumstances have a categorical reason—a reason that does not depend on what anyone wants or thinks—to prevent and condemn it.

B: On the contrary. In some circumstances abortion is morally required. Single teenage mothers with no resources have a categorical reason to abort.

C: You are both wrong. Abortion is never either morally required or morally forbidden. No one has categorical reason either way. It is always permissible and never mandatory, like cutting your fingernails.

D: You are all three wrong. Abortion is never either morally forbidden or morally required *or* morally permissible...A, B, and C are all assuming that moral duties exist. But no such thing exists, so none of them is making a true statement.<sup>49</sup>

While A, B, and C are making substantive moral claims about abortion, D is advocating Mackie's version of external error skepticism. Dworkin argues that even though D understands himself to be making a metaphysical claim *about* the moral domain, he is making a mistake about what his claim means, and thus what it commits him to. Dworkin writes:

...[H]e has seriously misunderstood the conversational situation. A, B, and C have each made a claim about what reasons of a certain kind—categorical reasons—people do or do not have. D's claim that no duties exist means that no one ever has a reason of that kind. So perforce he expresses a moral position. [...] What matters is not the arguments...but what they take to be the conclusion of those arguments. To repeat: each makes a claim about the categorical reasons people do or do not have with respect to abortion. The upshot of D's various arguments, whatever they are, is a claim of the same kind. He thinks there are no such reasons and therefore disagrees with A and B and agrees with C. He makes a much more general claim than C does, but his claim includes C's.<sup>50</sup>

Dworkin argues that his example demonstrates that purported external error skeptical claims can only be sensibly interpreted as abstract substantive claims because the former have the same *upshot* as do the latter.

I shall now argue that Dworkin has engineered this example in ways that prevent him from making a decisive case against his error skeptic opponents, which in turn prevents the example from supporting Substantive Stance. I pursue this conclusion by way of two objections.

---

<sup>49</sup> Dworkin (2011), p. 42.

<sup>50</sup> Dworkin (2011), p. 43.

### 3.2 Two Objections

The first objection holds that Dworkin illicitly incorporates content that his opponents will interpret as metaethical—presupposing MI—into the substantive claims made by A, B, and C, thereby preventing him from showing that (purportedly) metaethical claims—rejecting MI—are nothing but substantive claims. Dworkin imagines his discussants making claims about categorical reasons, which he characterizes as ones “that [do] not depend on what anyone wants or thinks.” This interpretation of what categorical reasons are makes accepting their existence tantamount to accepting MI. It should then be no surprise that D, who rejects MI, appears to make a claim that directly conflicts with those made by A, B, and C. But there are alternative ways of interpreting what categorical reasons are. For instance, one could offer different interpretations of the scope of those to whom such reasons apply, of the sense in which those reasons apply independently of what a person wants or believes, and of the practical meaning of claiming that there are categorical reasons to do or to avoid doing various things. Dworkin robs his argument of dialectical force by ignoring the possibility of these different interpretations. In so doing, he illicitly incorporates the MI component of the ordinary view into his example, when instead he needs to make the case for Substantive Stance *so that he can defend* the MI component.

The second objection to Dworkin’s use of the example holds that the unrealistic character of the “conversation” undermines its potential to support Dworkin’s conclusion. It is more than a little surprising how artificial Dworkin’s imagined dialogue is, given his emphasis on how ordinary people think about morality. People in substantive moral disagreements do not normally speak like A, B, and C do. Rather, when people find it necessary to spell out or offer further support for their moral stances, they normally do so by appealing to other, more general moral claims, which serve to elucidate their original claims

and, ideally, to find common ground with their interlocutors. They do not, in my experience, make appeals to the categoricity of their purported reasons.

Consider this reimagining of the conversation, which mirrors the structure of the original but involves more realistic content:<sup>51</sup>

A<sub>R</sub>: Abortion is always morally forbidden because fetuses are innocent persons. Intentionally killing an innocent person is murder, and murder is forbidden because it offends human dignity.

B<sub>R</sub>: You are mistaken. In some circumstances abortion is morally required, such as when giving birth would bring into the world a severely disabled child. In such a case, the suffering of the would-be child takes moral precedence. The way to respect human dignity in such a case is to prevent a very bad human life from being lived.

C<sub>R</sub>: You are both wrong. Abortion is never either morally required or morally forbidden, but always permissible. A<sub>R</sub> is wrong because a fetus is not a person, and so abortion is not murder. Therefore, a pregnant woman is not obligated to bring the pregnancy to term. B<sub>R</sub> is wrong because suffering is not as morally significant as B<sub>R</sub> claims; the future suffering of a would-be disabled child is not a sufficiently strong reason to forbid bringing a pregnancy to term.

The dialogue here remains idealized, and its language remains stilted in order to map onto Dworkin's original. Nevertheless, it presents a more realistic depiction of how people think and argue about moral issues. Notice that the dialogue depicts substantive moral debate as occurring at a certain remove from questions about whether MI is true. Consider how jarring it would be in this scenario if D were to interject, as in the original scenario, that abortion is never either morally forbidden, required, or permissible, because the world does not contain objective moral truths. A<sub>R</sub>, B<sub>R</sub>, and C<sub>R</sub> would likely respond with annoyance that they are trying to discuss what stance to take toward abortion, not to engage in abstract philosophical discussion.

In light of this fact, we should reconsider whether it makes sense to conceive of D as taking part in the discussion at all. Recall Dworkin's complaint that D "has seriously misunderstood the conversational situation." This claim is a bit rich, of course, since *it is*

---

<sup>51</sup> The 'r' below stands for "realistic."

*Dworkin who has placed him there.* If D were trying to advocate external error skepticism, he would probably not do so by engaging others in a substantive moral debate. Here are A, B, and C, engaged in substantive discussion about what there is moral reason to do, and along comes D: “Oh, you believe there are objective moral reasons? Haha, big mistake, you dummies!”

The general problem here is that Dworkin’s example posits a highly implausible context, and *conversational context affects conversational content*. In order to understand what someone is saying, we often need to know the context in which he is saying it, including what he *intends* to be saying. Dworkin strips away this latter element of conversational context by suggesting that D would state (what D takes to be) his metaethical view out of the blue in an ongoing substantive discussion. A number of error theorists—including Mackie himself—have engaged in substantive moral discourse while also holding the (purportedly metaethical) view that MI is false, so it seems reasonable to consider the context of the error theorist’s utterance before determining whether he is making a substantive or a metaethical claim.<sup>52</sup>

In addition, note that Dworkin leaves out of D’s statement information that would be crucial to A, B, and C understanding what D is saying. Since D is advocating error theory, it is important to include his rationale: categorical duties would have to be objectively prescriptive to exist, but since there is nothing objectively prescriptive, there are no categorical duties. If D were to include this information, it would surely tip off A, B, and C that he was addressing a different issue than what stance to take toward abortion. Moreover, it would allow A, B, or C the possibility of disputing the foundation of D’s view—the claim that categorical duties would have to be objectively prescriptive to exist—before returning to

---

<sup>52</sup> I say more about this possibility in §3.3 below.

the substantive disagreement with everyone on the same page, or at least with a better understanding of what each disputant has in mind. Dworkin does not consider that A, B, or C might disagree with D that MI would have to be true in order for their substantive moral discussion to be justified, likely because that very presumption is built into the ordinary view. While Dworkin claims that his view is widely held, we have already discussed why this claim is almost certainly false, so he should not build into his example that all three other disputants happen to share the presumption of MI contained in that view.<sup>53</sup> We might imagine A replying to D's stated rationale by saying, "Oh, I'm not trying to take a stand on whether the universe contains objectively prescriptive norms. I'm just arguing that abortion is morally wrong."

### 3.3 Responses & Further Objections

Dworkin might offer two responses to the above objections. First, maybe we *can* view D as a rude conversational interloper who makes his skeptical claim with the aim of pulling the rug out from under A, B, and C, whose moral claims he views as fundamentally misguided precisely because he thinks MI is false. Second, even if D were a polite conversationalist, what matters is not what D intends to be claiming, but what D is in fact claiming. Dworkin maintains, after all, that philosophers who take themselves to be doing metaethics are simply misunderstanding what they are up to, so he would be comfortable claiming that D is just mistaken about the nature of what he is claiming. Recall the second half of the passage quoted above:

What matters is not the arguments...but what they take to be the conclusion of those arguments. To repeat: each makes a claim about the categorical reasons people do or do not have with respect to abortion. The upshot of D's various arguments, whatever they are, is a claim of the same kind. He thinks there are no such reasons

---

<sup>53</sup> Note that this point shares a similar spirit with the first objection.

and therefore disagrees with A and B and agrees with C. He makes a much more general claim than C does, but his claim includes C's.

If we again consider the original scenario—setting aside for now the worries about built-in appeals to categorical reasons—Dworkin's idea is that the “upshot” of D's claim that there are no categorical reasons of any kind places him in the same boat as C, who claims that abortion is always permissible, never obligatory or forbidden.

It is not immediately clear how Dworkin understands the concept of a claim's upshot. It is not quite right to say, as Dworkin does, that D's “claim includes C's,” since D rejects as erroneous *all* substantive moral concepts, including that of permissibility.<sup>54</sup> However, there is another view in the vicinity: given the claims of C and D, for all *practical* purposes D should behave the same as C with regard to abortion. D should act as though abortion were permissible, since he should act as though *everything* were permissible, since he believes there are no moral constraints of any kind. Dworkin provides further evidence for this interpretation of his view when he claims that both internal skepticism and external error skepticism “play for keeps,” by which he means that they “[have] direct implications for action.”<sup>55</sup> He elaborates:

If someone is internally skeptical about sexual morality, he cannot consistently censure people for their sexual choices or lobby for outlawing homosexuality on moral grounds... [Likewise,] an error skeptic may dislike the war in Iraq, but he cannot claim that the American invasion was immoral.”<sup>56</sup>

Dworkin's use of the phrase “cannot consistently” in this passage makes it clear that his position is not just that both forms of skepticism have direct implications *for action*, but that both have direct implications *for the assessment of action*.<sup>57</sup> To say that a person is acting

---

<sup>54</sup> The concept of permissibility presupposes that there are moral rules *allowing* certain actions.

<sup>55</sup> *Ibid.*, p. 35.

<sup>56</sup> *Ibid.*, pp. 35-6.

<sup>57</sup> Here, action must be interpreted broadly so that it includes the making of claims.

inconsistently with her professed moral belief is to suggest that she is acting irrationally or hypocritically, or else that she does not actually hold the moral belief she professes to hold.

It is a familiar point that skeptics cannot “live their skepticism.” The external world skeptic still feeds the cat and tries to dodge the oncoming boulder, and the moral skeptic still resents being lied to or stolen from. While commentators on both sides of the debate often treat these facts merely as manifestations of practical necessities or ineliminable dispositions, or as balms for worries about the practical effects of the widespread adoption of skepticism, Dworkin treats them as sources of rational pressure on the skeptic to abandon his view.

Here is how he characterizes the situation in an earlier article:

...[A]ny reason we think we have for abandoning a conviction is itself just another conviction, and...we can do no better for any claim, including the most sophisticated skeptical argument or thesis, than to see whether, after the best thought we find appropriate, we think it so. If you can't help believing something, steadily and wholeheartedly, you'd better believe it. Not, as I just said, because the fact of your belief argues for its own truth, but because you cannot think any argument a decisive refutation of a belief it does not even dent. In the beginning, and in the end, is the conviction.<sup>58</sup>

Here, Dworkin relies on Substantive Stance to claim that if D still finds himself believing substantive moral claims in practice, despite his (purportedly) theoretical acceptance of external error skepticism, D should drop the error skepticism, since he clearly does not really believe it.

There are two problems with this analysis. First, of course, since it relies on Substantive Stance, we cannot use it to support Substantive Stance. Second, it is not clear how we should understand the normative character of the “you'd better” being asserted. Is it a moral principle or is it a general epistemic principle? Since Dworkin is recommending a particular way of adjudicating between purportedly moral claims, it might seem plausible that he is advocating a moral principle. However, we cannot expect D to accept it, since he

---

<sup>58</sup> Dworkin (1996), p. 118.

denies that there are any valid moral principles. If it is an epistemic principle, on the other hand, it is simply implausible. Just because I cannot help but act as though some belief, *x*, is true does not necessarily mean I should abandon my belief that *x* is false.

Imagine that I deny the existence of ghosts. Nevertheless, when I visit a house that is reportedly haunted, I find myself jumping at every sound, seeming to see strange things out of the corner of my eye, and feeling as though my movements are being tracked by an unseen presence. Does that mean that I actually *do* believe in ghosts? Maybe it does. After all, I am *acting like* I believe there is a ghost in the house. On the other hand, maybe my reactions in the house are, for whatever reason, not responsive to my beliefs about what kinds of entities the world contains, or about which kinds of situations warrant fear. Suppose this latter interpretation is more plausible, and that it means I am behaving irrationally. Would the irrationality of my state give me reason to accept the existence of ghosts? Presumably it would not. Even if I am open to rational criticism when there is a conflict between my beliefs and my reactions, there is no general norm that instructs me which of the two—my beliefs or my reactions—I must revise in order to be acting rationally.

Now suppose we set these particular worries aside and grant Dworkin that the error skeptic's belief *does* have the practical upshot he proposes. After all, it does seem reasonable to claim that the error theorist is mistaken in thinking he can consistently advocate both error theory and substantive moral views, since his theory holds that all substantive moral views are false. If one claims that there are no *Xs*, it would be strange if one then started referring to *Xs*, absent some account that explained why this move is permissible. Nevertheless, even if we admitted that advocating error theory has this practical upshot, it would not establish Substantive Stance. It might just be that some beliefs in external claims can have practical upshots for beliefs in internal claims. In other words, one could maintain

that there is a genuine distinction between metaethical and substantive claims, but maintain that that some metaethical views influence which, if any, substantive views one can consistently hold. There is a significant difference between *that* claim and the claim that there is *no such thing* as a distinctively metaethical view, that all purportedly metaethical beliefs *are nothing more than* substantive beliefs. This is a big problem for Dworkin, since error theory is the most extreme form of (purportedly) external moral skepticism, and so would seem to present his best opportunity for establishing Substantive Stance.

By contrast, consider how Dworkin's approach would attempt to deal with the skeptical challenge from Chapter 1, which holds that we lack grounds for rational confidence in MI, but which makes no claims about whether MI is true. How plausible is it to claim that the skeptical challenge is nothing more than a substantive moral position? A standard way of understanding the view, of course, would treat it as a metaethical epistemic claim, and so as orthogonal to substantive moral claims. But Dworkin is committed to treating all forms of skepticism about morality as abstract moral claims. Thus, he would treat a defender of the skeptical challenge as a fifth discussant in the conversation about abortion:

E: We lack grounds for rational confidence in the view that morality is objective.

How, if at all, does E's claim interact with the substantive claims about abortion made by A, B, and C? What is its practical upshot?

In order to defend Substantive Stance, Dworkin would have to show that E's claim has the same practical upshot as a regular substantive claim. I suspect he would treat E's claim as *rejecting* every argument in favor of a particular moral claim, perhaps under the rationale that if one believes there are no grounds for rational confidence in some claim, one should not believe that claim. In this case, then, E would reject every claim that labels abortion obligatory, forbidden, or permissible, which would place E in direct conflict with A,

B, and C (or, as Dworkin would have it, in conflict with A and B but in agreement with C). For all practical purposes, then, E would be indistinguishable from the error theorist—she would think people who endorse moral claims are making a mistake, and she could not consistently endorse moral claims herself. Dworkin would thus object to (this interpretation of) the skeptical challenge just like he objects to error theory. Both theories would be self-defeating and would have upshots that conflict with the upshots of A's, B's, and C's more plausible substantive claims.<sup>59</sup>

Two major assumptions underlie this Dworkinian treatment of the skeptical challenge. The first assumption holds that claiming that a belief lacks some positive epistemic status—in this case, grounds for rational confidence—has the same practical upshot as does claiming that the belief is false, and thus has the same practical upshot as does claiming that one should not hold that belief. The second assumption holds that a person could not consistently believe particular moral claims while not believing MI. There are serious problems with both assumptions. The first assumption rides roughshod over the important difference between challenging a belief's epistemic status and labeling it false. In the present case, Dworkin needs to show that both kinds of challenge have moral upshots indistinguishable from those of substantive claims, but he would need *first* to establish Substantive Stance for this move to be legitimate. Yet again, he is in the unenviable position of relying upon the very claim he needs to prove. A similar problem plagues the second assumption: it presumes a tenet of the ordinary view rather than arguing for it. It presumes that holding substantive moral beliefs commits one to MI. This presumption has also been at work in Dworkin's discussion of error theory, but is less dialectically problematic in that context because error theory shares that presumption. The skeptical challenge does not share

---

<sup>59</sup> The skeptical challenge would be self-defeating because it would itself be a substantive moral claim that denies we have grounds for rational confidence in believing that very claim.

it, however, leaving open the possibility that E could explicitly *agree with* A, B, or C as a substantive matter, while still advocating her skeptical challenge, as follows:

Expansive E: In some circumstances, abortion is morally required. Nevertheless, we lack grounds for rational confidence in the view that morality is objective. So we must understand the nature of that moral requirement in a way that does not commit us to MI.

Of course, for all the reasons given in §3.2, it would be unrealistic for Expansive E to bring up both of these views in the same conversation.

The upshot of this discussion is that Dworkin cannot legitimately treat E's claim as equivalent to a claim that there are no moral truths—including truths about abortion—since doing so would require presuming two tenets of the ordinary view that he is in the process of arguing for. I believe that these objections undermine the arguments for Substantive Stance that Dworkin has so far offered. However, two significant issues remain. The first is whether Dworkin's arguments against status skepticism are more effective than are his arguments against error skepticism. The second is whether there is a good case to be made for the second assumption just discussed: the idea that holding substantive moral beliefs commits one to MI. These two issues are closely linked, as we will see in the next section.

## **4. Dworkin on Status Skepticism**

### **4.1 The View and Its Apparent Problems**

Status skepticism is the other form of (purported) external skepticism that Dworkin considers. This view rejects MI by holding that moral claims are not apt for objective truth or falsity, since they play some function other than to describe how things are, morally speaking. Dworkin has in mind such views as emotivism, prescriptivism, expressivism, and quasi-realism. As I explained in §2.2, it may be sensible to include under the heading of status skepticism other (purportedly) metaethical views that treat moral claims as capable of

*non-objective* truth. Indeed, it will become clearer over the next several paragraphs that Dworkin's main issue is not whether moral truths are in the business of *describing* moral reality, but instead whether we have good reason to understand moral reality as objective at all. Thus, I will treat both the skeptical challenge and Intersubjectivism as forms of status skepticism, which means my arguments against Dworkin's treatment of that view are therefore also arguments for the viability of my own views.

Dworkin explains that status skepticism has attracted proponents "because it does not ask us to pretend we are abandoning convictions that we cannot actually abandon. It encourages us to keep our convictions and give up only bad metaphysics."<sup>60</sup> The idea motivating the view is that revising our understanding of what moral commitments are or involve will allow us to avoid the metaphysical and epistemological complications that beset MI, and thereby to engage in moral discourse with a cleaner philosophical conscience. Of course, Dworkin believes that only *moral* conscience is relevant to assessing the status of morality, since any claim that could affect its status would necessarily be a moral claim. But we, as readers, are still waiting for a good argument for this view.

Unfortunately, Dworkin's approach to attacking status skepticism is almost identical to his approach to attacking error skepticism. He once again offers examples that he believes show there is no genuine difference between substantive moral claims and purportedly metaethical claims. Dworkin writes,

[Status skepticism] is available, even as a position to contest, only if we can establish a distinction between what the two following judgments mean or come to: first, that torture is always wrong, and second, that the wrongness of torture is a matter of objective truth that does not depend on anyone's attitudes. If the second, supposedly philosophical, judgment is only a wordy restatement of the first concededly moral one, then no one can coherently embrace the first without the second and status skepticism is a bust from the start...It doesn't help to insist, as many status skeptics do, that the first-order judgment that torture is wrong is only the projection of an

---

<sup>60</sup> Ibid., p. 52.

attitude and not really a judgment at all. If it is, then why isn't status skepticism just the projection of the opposite attitude and not a philosophical position at all?...I believe [this] challenge is fatal to all forms of that view.<sup>61</sup>

The status skeptic wants to be able to license (in different ways for different theories) the first, substantive claim about the wrongness of torture while denying the second, purportedly metaethical claim about the objectivity of that wrongness. If she cannot justify this maneuver, then she cannot state her skeptical theory.

Justifying the maneuver requires her to show that there is or can be a meaningful difference between these two claims, such that they do not stand or fall together:

**Moral Claim:** Torture is always wrong.

**Further Claim:** The wrongness of torture is a matter of objective truth that does not depend on anyone's attitudes.<sup>62</sup>

Dworkin asserts that Further Claim is nothing more than a "wordy restatement" of Moral Claim, such that there is never a meaningful difference between the two. If the status skeptic endorses Moral Claim, then she contradicts herself in denying Further Claim. Or, to pose the problem in more relevant terms, if she rejects Further Claim, as she must to articulate status skepticism, then she contradicts herself in accepting Moral Claim. Thus, Dworkin relies on Substantive Stance to argue that one cannot state status skepticism without contradicting the very moral commitments one is presumably trying to protect.<sup>63 64</sup>

Let's now consider whether Dworkin has good reason to rely on Substantive Stance in this way. The only evidence he marshals in favor of this move is the purported equality in

---

<sup>61</sup> Ibid.

<sup>62</sup> I am following Dworkin in using this "further claim" terminology.

<sup>63</sup> Status skeptics need not understand themselves as being in the business of trying to protect or license morality. Some may merely be offering a view of how morality—whatever its status turns out to be—fits into the world.

<sup>64</sup> Notice the parallel with Dworkin's treatment of error skepticism, where he relies on Substantive Stance to argue that error skepticism is self-defeating because it applies to itself. In both cases, he argues that (purportedly) external skepticism is not just mistaken, but impossible even to maintain.

meaning or upshot of Moral Claim and Further Claim.<sup>65</sup> Does claiming that some act is always wrong necessarily imply a commitment to MI? I have already argued that it does not. I have appealed to considerations that are equally relevant in the present case, such as the context sensitivity of meaning and the alternative ways that categorical statements can be interpreted, so I will not revisit that territory here. Instead, I'd like to explore a distinct approach Dworkin could take to interpreting what it means for a claim to have a certain upshot. This approach might allow him to motivate acceptance of Substantive Stance in the face of status skepticism by showing that Moral Claim and Further Claim have the same upshot.

## 4.2 Two Kinds of Upshot

In §3.3, I explained that the most plausible way to interpret Dworkin's idea of a claim's upshot is in practical terms: how one's belief in some claim should practically affect how one acts. And the primary concept I considered in assessing a claim's upshot was that of acting consistently with the claim's (purported) semantic implications. Let's say that we can use this concept to identify a claim's *semantic upshot*. There is a second concept that may be just as important to assessing a claim's upshot—that of what is psychologically possible for a believer. Employing this concept, we might interpret Dworkin's worry about status

---

<sup>65</sup> I say this is the only evidence because he uses the same pattern of argument repeatedly. For instance, here is Dworkin using the example of abortion:

I am speaking at length about abortion. I begin: "Abortion is morally wrong." Then, drawing breath, I add a variety of other claims set out in the rest of this paragraph. "What I just said about abortion was not just venting my emotions or describing or expressing or projecting my own or anyone else's attitudes or my own or anyone else's commitment to rules or plans. My claims about the immorality of abortion are really, objectively, true. They describe what morality, quite apart from anyone's impulses and emotions, really demands." [...] Call all the statements I made after drawing breath my "further claim." [...] My further claims also appear themselves to be moral claims. If so, and [the status skeptic] denies them, he makes a moral claim as well. (Ibid., p. 53.)

skepticism as arising from the idea that it is not psychologically possible for a person to be skeptical about the status of morality and yet still go on to make categorical moral claims. Let's say that we can use this concept to identify a claim's *psychological upshot*. We could then interpret Dworkin's objection as holding that one's endorsement of status skepticism could not remain merely theoretical, but would necessarily infiltrate one's practical life, and would therefore carry with it the same psychological upshot as would endorsing global moral skepticism. Put differently, the objection would hold that defending status skepticism requires positing an unrealistic division in an agent's mental life.

There appear to be several virtues of employing the concept of a psychological upshot. It would help make sense of why Dworkin embeds a commitment to MI within the ordinary view, since that view is meant to characterize what people actually believe, and, by extension, to rule out what people *could not* believe. It would also add a much-needed rationale for embracing Substantive Stance, in light of the weakness of the rationales so far offered. Moreover, endorsing the concept of a psychological upshot is consistent with the motivation for embracing MI that I posited when discussing practical BS-ing in Chapter 1. I hypothesized there that some defenders of MI worry that rejecting the view would prevent one from being able to take everyday moral thought and discourse seriously. This seems to be the exact worry to which Dworkin is giving voice.<sup>66</sup>

The final virtue can be seen by analogy with other situations in which being skeptical about the status of some domain might conflict with a person's ability to operate within that domain in the normal way. Suppose that Jim has long believed that finding true love is a matter of meeting the unique person chosen by God to be one's life-long companion. For

---

<sup>66</sup> Dworkin differs from other proponents of MI, however, in that he denies that one could fail to take moral thought seriously, and so denies that one could actually endorse external skepticism. He thinks we are all necessarily proponents of MI, even though some of us do not realize it.

many years, this belief has shaped how Jim understands his place in the world and his interactions with potential mates. However, Jim has recently read some books about evolutionary psychology, which offer a starkly different depiction of the forces shaping human life. He has now come to believe that what he once called “true love” is actually nothing more than a chemically induced illusion conjured by our genes to improve our reproductive odds. As a result, Jim now counts himself a skeptic about true love.

It seems reasonable to suppose that Jim’s newfound skepticism will exert a psychological influence on how he interacts with potential mates. Every time he feels romantic affection welling up in his breast, perhaps he will bitterly remind himself that we are no more than puppets cast in a biological play, our actions directed by the thoughtless masters of competition and chance. This will ruin Jim’s dates. It may seem reasonable to suppose that status skepticism about morality would have similar psychological effects on its proponents. If the marauder feels a pang of conscience before she touches her torch to the hut, recalling that morality is not apt for objective truth may be all it takes for her to gleefully send its straw roof up in flames.

#### **4.3 Doubts about Psychological Upshots**

Despite the apparent virtues of this approach, I shall now argue that appealing to the psychological upshot of endorsing status skepticism does not help Dworkin establish Substantive Stance. The crux of the issue is whether being skeptical that moral claims are apt for objective truth necessarily undermines one’s disposition or ability to engage in substantive moral discourse. Luckily, we can compare this case to those of several other domains that share two features: we are skeptical that their governing norms are objectively true, yet we are still able to engage with those norms substantively. Consider the domain of

grammar. It is uncontroversial that grammatical norms are intersubjective, not objective. These norms have been cobbled together and refined in different ways, for different languages and at different times, all for the purposes of promoting communication, enabling social signaling, and so on. We might say, then, that most reflective people are status skeptics about the objectivity of grammatical norms. Nevertheless, we have no problem engaging in substantive discussion about which grammatical norms to follow in which contexts, and we often make grammatical claims that are categorical in form. In short, our status skepticism about the objectivity of grammatical norms does not present a psychological obstacle to our taking those norms seriously in practice.

We can illustrate this point using an example parallel to Dworkin's. Consider these two claims:

**Comma Claim:** It is always wrong to use an Oxford comma.

**Further Comma Claim:** The wrongness of using an Oxford comma is a matter of objective truth that does not depend on anyone's attitudes.

Further Comma Claim is not merely a "wordy restatement" of Comma Claim, even though the latter is categorical in form. Moreover, it would be ludicrous to hold that rejecting Further Comma Claim would make it psychologically impossible to accept Comma Claim. This discussion raises the following challenge: since one can be skeptical that some normative domains deal in objective truths without thereby precluding oneself from making categorical substantive claims within those domains, Dworkin owes us a special reason to think that morality operates differently.

Suppose that Dworkin would object that this analogy is inapt, since it is a *moral* issue whether grammatical norms are objectively binding. Rules for or against using an Oxford comma lack the kinds of normative weight or application possessed by rules for or against, say, torture. Nearly everyone would agree with this last claim as a substantive matter, of

course. But recall that we are tackling a different issue: the psychological upshot of denying the objectivity of some normative claim. Given where we are in the dialectic, Dworkin could not simply *assert* that this upshot is different in the moral case than in the grammatical case. That is the very issue in question.

We can also examine this dynamic from the opposite direction by considering what, if anything, one's making of categorical claims within a normative domain implies about one's commitment to the objectivity of that domain. Recall that one of my objections to Dworkin's treatment of the abortion discussion was that he builds into his scenario an interpretation of categorical claims that presupposes MI. We can see further evidence of why this move is illicit by considering a parallel scenario in which people are discussing grammatical rules:

A<sub>C</sub>: You must always use an Oxford comma between the last two items in a list of three or more, since doing so helps to avoid ambiguity.

B<sub>C</sub>: I disagree. You should never use an Oxford comma, since it is better to avoid ambiguity by rephrasing the problematic sentence.

C<sub>C</sub>: You are both wrong. An Oxford comma is never either obligatory or forbidden. Using it is a matter of personal preference, so is always merely permissible.

D<sub>C</sub>: You are all three wrong. An Oxford comma is never either obligatory or forbidden *or* permissible, because the world does not contain objective grammatical norms. Rather, those norms are merely intersubjective.

Although the rules of grammar make claims about how a speaker ought or ought not construct a sentence, they do not imply any claims about the objectivity of this "grammatical ought." The challenge for any defender of Substantive Stance is to show whether and, if so, how embracing a categorical moral commitment functions differently from embracing a categorical commitment in another normative domain. Dworkin believes that the "moral ought" carries with it a claim to objectivity, but it is difficult to see how appealing to that notion at this point would be anything other than begging the question in favor of his

ordinary view, which takes for granted exactly what is at issue, i.e., whether engaging in moral discourse commits one to MI.

We can make a similar point with regard to the rules of games. For instance, the rules of basketball hold that a player in possession of the ball must not take more than two steps without dribbling. To transverse this rule is to commit a traveling violation, punishable by loss of possession. Players take this rule seriously during a game. They use it to guide their actions, they hold one another accountable to it, and so on. But taking the “no traveling” rule seriously does not commit players to thinking that the normativity of the rule is objective. Any question about the rule’s metanormative status is altogether separate. And the answer to that question, as most people realize, is that the rule is intersubjective, not objective. Just as with linguistic rules, the rules of games have been created and refined by people in order to aid our joint activities, and to enable us to produce and enjoy things we care about.

I have undertaken two main tasks in this section. First, given that Dworkin’s arguments for Substantive Stance and against status skepticism are subject to the same vulnerabilities I discussed in §3, I have proposed a novel way of understanding a claim’s upshot that focuses on psychological possibility. Second, I have considered whether appealing to the notion of a psychological upshot allows Dworkin to establish Substantive Stance and thereby to defeat status skepticism. I have argued that it does not. We have good evidence from other normative domains that one can harbor skepticism about a domain’s objectivity without thereby precluding oneself from engaging with its norms substantively. This phenomenon is not merely psychologically *possible*; it is utterly commonplace.

## 5. Conclusion

I believe that this discussion defeats Dworkin's attempts to establish Substantive Stance and, by extension, his attack on the viability of external skepticism. It preserves the skeptical challenge from Chapter 1 and sets the stage for the arguments in favor of Intersubjectivism in Chapter 3. Nevertheless, our discussion of Dworkin's view has raised important questions about the psychological basis of our moral interactions, and about the ways in which morality may be importantly different from other normative domains. I pick up on these questions to begin the next chapter.

## INTERSUBJECTIVISM: A THEORY OF MORAL NORMATIVITY

### 1. Introduction

My aim in this chapter is to show that Intersubjectivism is a promising candidate for solving the reconciliation problem. I begin in §2 by clarifying the nature of that problem and outlining my strategy for addressing it. Next, I advance an account of human psychology that lays the groundwork for understanding moral normativity as essentially intersubjective. In §3, I discuss how Modern thinkers sought to explain what is most distinctive about human social life by appealing to our *drive for recognition*. I pick up on this tradition in §4 by arguing that one species of this drive—the *drive for respect*—enables us to participate in moral relationships and thereby helps to make us persons. In §5, I draw on this account to argue that we together create and maintain moral normativity through our activities of granting one another practical authority within moral relationships.

### 2. A Strategy for Reconciliation

#### 2.1 Lingerin g Concerns

I argued in Chapter 2 that a person's skepticism about moral objectivity need not undermine her disposition or ability to engage in substantive moral discourse. I made the case for this claim, in part, by drawing analogies with our abilities to participate in other norm-governed activities, such as writing grammatically or playing a game. However, if these analogies are apt, why is there a vast amount of writing defending moral objectivity but virtually none defending the objectivity of grammatical rules or the rules of basketball? Why would so many philosophers happily admit that the latter two are intersubjective, but hold

that the former must be objective if they are to be philosophically respectable? The answer, I contend, is that many people conceive of morality in such a way that anything less than an objective foundation for our moral commitments would seem too flimsy to support their normative weight. Given this conception, an intersubjective foundation would seem to be no foundation at all.

It is easy to understand this worry when we consider our attitudes toward the norms that underlie our moral commitments. Three features stand out. First, we treat these norms as *overwhelmingly important*. Our moral commitments shape how we understand and organize our lives because, taken together, they express a conception of who we are, what we care about, and how we relate to others. Moreover, whether others share or respect our moral commitments has an enormous effect on how well our lives go. Our beliefs about the rules of games and grammar, by contrast, lack this kind of overwhelming importance.

Second, we take the content of our moral commitments to be *non-contingent*. We could easily change the rules of basketball by agreeing that a player is allowed three steps without dribbling instead of just two, but it seems clear that pillaging would still be wrong even if we all thought it was permissible. Put differently, we must leave conceptual space between moral normativity and whatever moral beliefs people happen to hold, since we are liable to get things wrong, even, or sometimes especially, when we are thinking and acting together.

Third, we take the content of our moral commitments to be *authoritative*. We often make moral claims on others in contexts of interpersonal disagreement about how people should act. If disagreement is the source of the problem—the scene and the catalyst of our making those moral claims—it is difficult to see how positing an intersubjective foundation for morality could be of any help. The *very problem* in these contexts seems to be that we lack

intersubjective agreement. If moral disagreement is to involve anything beyond a brute clash of attitudes or interests, it seems we need the moral norms that underlie our claims to serve as authoritative arbiters of disputes. And they can only play this role, we might think, if they are independent of the attitudes people happen to have, even if those attitudes happen to be in fortunate agreement.

In sum, the overwhelming importance, non-contingency, and authority of moral norms seem to distinguish them from the rules of games or grammar. It may seem reasonable to conclude that the source of these significant differences is the mere intersubjectivity of the latter domains, and therefore also to conclude that an intersubjective basis would be too rickety to support a normative domain with morality's distinctive features.

Do these observations show that I was too hasty in dismissing the claim that a person's rejection of MI has a detrimental psychological upshot for the way she engages in moral thought? They may seem to. After all, if a person's conception of a normative domain could not make sense of features essential to her engagement in that domain, we would expect her to be unable to engage in it in the normal way, unless her psychological life were so poorly integrated that her conception of what she was doing diverged wildly from what she actually was doing. But recall that the question about psychological upshots asked not whether rejecting MI *could* have a detrimental effect on a person's moral behavior—that certainly seems possible—but whether it *necessarily would*. We can answer this latter question in the affirmative only by showing that no MD conception could successfully account for morality's essential features. Therefore, I suggest that we treat the above observations as guides to those features of morality that Intersubjectivism will need either to explain or to explain away.

We can bring out the general challenge posed by psychological upshots by again considering Jim. His recent study of evolutionary psychology has led him to become a skeptic about true love, and that skepticism has made it difficult for him to engage in normal romantic relationships. But Jim's main problem is not that he has lost faith in his old conception of true love. His problem is that he has not yet replaced it with a conception capable of reconciling his experience of love with his newfound understanding of the biological and cultural forces shaping that experience. So Jim faces a reconciliation problem of his own. He can only resolve that problem by adopting a conception of love that coheres with his naturalistic conception while also doing justice to his practical, lived experience. For example, he could conceive of love in terms of one's wanting to share central aspects of one's life with another, valuing the other for the other's own sake, practically committing to care for the other, and so on. This conception employs concepts—such as desiring shared experience, valuing, caring, and practically committing—that are at once integral to the experience of love and compatible with any plausible causal explanation of what we are doing when we love. It can be true both that Jim's genes strongly dispose him to develop feelings of love and that those feelings draw him into actual loving relationships. If Jim were to adopt this conception, he would no longer be skeptical that love exists. He would just have a new, more plausible understanding of what its existence amounts to.

In this chapter, I pursue a parallel approach to arguing that Intersubjectivism allows us to solve the reconciliation problem for morality. I argue that we can explain moral normativity in terms of our practical commitment to participating in relationships of mutual respect with other persons. This explanation coheres with a plausible naturalistic account of why humans are moral creatures, yet does so by employing concepts—such as personhood,

practical commitment, mutual respect, and accountability—that are integral to moral thought. Or so I aim to show.

## 2.2 Reconciliation Redux

I have argued that the reconciliation problem arises because it is difficult to reconcile our practical and naturalistic conceptions of our moral lives, yet we can abandon neither. While the reconciliation problem clearly presents us with a philosophical challenge, it can also affect us practically. As rational and self-conscious beings, we often guide our actions by employing a practical conception of what we are doing. This capacity enables us to act in sophisticated ways, but it also makes us vulnerable when the process becomes destabilized. As the case of Jim demonstrates, losing confidence in a practical conception can prevent us from acting as we normally would. Periods of depression or existential crises confront people with similar challenges. The lethargy characteristic of depression follows from a person's losing her grasp on the practical conception under which her life and activities have value. I don't mean to suggest that metaethical inquiry has the same stakes, of course. But I have proposed that being aware of the stakes—however we choose to characterize them—plays a role in some theorists' decisions to defend MI.

I draw attention to this dimension of the reconciliation problem in order to show that solving it would be an achievement that is both philosophical *and* psychological. As with Jim's grappling with the meaning of romantic love, we grapple with big metaethical questions not as disinterested observers, but rather as moral beings trying to make sense of how our moral commitments fit within our broader conception of what we and the world are like. Insofar as we are unable to reconcile our practical and naturalistic conceptions of our moral lives, our practical moral engagement may feel unmoored. The concept of having

*grounds for rational confidence* in a metaethical view—introduced in Chapter 1—reflects the dual nature of the reconciliation problem. The notion of having rational grounds captures the aspect of epistemic appraisal, while the notion of having confidence captures the way in which taking a view to be successful would, or could, affect a person’s psychological state.

What would it take to solve the reconciliation problem and thereby generate grounds for rational confidence in a metaethical view? There may be several plausible approaches. The approach I pursue identifies the source of the reconciliation problem—the *tension* between two fundamentally different ways of conceiving of ourselves as moral beings—and tailors the solution to address it. Suppose you are holding a piece of string between two hands. You can create tension in the string by moving your hands farther apart, and you can relieve tension by bringing them closer together. Likewise, my attempt to relieve the tension between our practical and naturalistic conceptions offers an account that brings the two closer together. Intersubjectivism defends an account of moral normativity that aims to capture and clarify our practical conception, while cohering with a naturalistic conception that does justice to the psychological dynamics of moral relationships.

### 2.3 Intersubjectivism and Its Aims

One of the primary aims of metaethics is to tell us *what we are doing* when we take ourselves to have moral reasons, or when we claim that particular actions are obligatory, permissible, or forbidden. One way to pursue this aim is to explain how moral concepts function and what roles they play in our lives.<sup>67</sup> These explanations gain plausibility to the

---

<sup>67</sup> Note that there are different concepts of concepts, as it were, and not all of them treat concepts as having functions.

extent that they also help us understand where moral concepts come from and what we need them for.<sup>68</sup> In this chapter, I argue for

**Intersubjectivism:**

Paradigmatically, moral claims are appeals to practical authority issued within relationships of mutual respect and accountability. Moral normativity arises from our activities of granting one another practical authority within these relationships.

The first component of the view characterizes the function of moral concepts and the contexts in which we use them. The second component holds that we can understand moral normativity as intersubjective, something we create and maintain together by participating in moral relationships.<sup>69</sup>

Let me now offer some clarificatory remarks about each part of the view. The reader will notice that my statement of Intersubjectivism begins, somewhat inauspiciously, with the hedge term “paradigmatically.” I believe that the dynamics of making moral claims are best understood in terms of persons holding one another accountable to the standards that structure their interpersonal relationships. In making moral claims, then, we necessarily express a conception of how we relate to each other.<sup>70</sup> Nevertheless, interpersonal morality does not exhaust the moral domain. We may also take ourselves to have moral duties to

---

<sup>68</sup> Different kinds of theories will face these explanatory burdens and/or opportunities to different degrees. Most secular MI theories will have less to say about these last two issues: our moral concepts correspond (if we are lucky) to moral truths that are independent of us, and we need these concepts in order to ensure that we act in accordance with those independent moral truths. MD theories have more that they need to say about these matters (and so a greater explanatory burden), but also more that they *can* say about them (and so greater explanatory potential).

<sup>69</sup> One could endorse the first component without also endorsing the second. I doubt that one could endorse the second without also endorsing the first.

<sup>70</sup> Compare this claim to Christine M. Korsgaard’s contention that

The primal scene of morality...is not one in which I do something to you or you do something to me, but one in which we do something together. The subject matter of morality is not what we should bring about, but how we should relate to one another. (“The Reasons We Can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values,” *Social Philosophy & Policy*, Vol. 10, No. 1 (1993), pp. 24-25.) It is natural to read Korsgaard as making as a substantive moral claim here, with the view I defend being a nearby metaethical counterpart. However, since she attempts to derive the substantive principles of morality from the formal requirements of agency, her claim does not fit straightforwardly within the substantive-metaethical dichotomy.

beings that are not persons, such as non-human animals, young children, and non-sentient aspects of the natural world. In addition, we may take ourselves to have moral duties to help bring about states of the world, such a future that does not contain miserable hordes facing environmental collapse.<sup>71</sup>

My strategy is to focus on understanding the paradigmatic, interpersonal context, setting aside for now these other moral contexts. I pursue this strategy for three reasons. First, interpersonal morality is the domain about which critics are most likely to think an MD theory is objectionably revisionary, given the immense practical importance of people taking themselves to be bound by others' claims. Second, one of my goals is to shed light on where our moral concepts come from, and our moral thinking plausibly evolved to deal with interpersonal contexts before later being extended to other contexts.<sup>72</sup> Third, it is possible that whatever explains the normativity of interpersonal morality cannot also explain the normativity of other parts of morality. In the Introduction, I noted my skepticism that we can articulate a single, unified account of normativity capable of producing illuminating explanations of every normative domain. Something similar could be true *within* the moral domain.<sup>73</sup> If there is no unified theory of moral normativity to be had, then holding out for such a theory can prevent us from finding success locally. On the other hand, if there is such a unified theory, local success can only help us make progress toward it. For now, then, my focus will be on the unity and structure of interpersonal morality.

---

<sup>71</sup> Though it is famously difficult to articulate the content of these duties, as Derek Parfit has shown. See Parfit, *Reasons and Persons* (New York: Oxford University Press, 1984), Part 4.

<sup>72</sup> Compare Korsgaard's claim that "deontological restrictions predate [the] global issues [of population control or the preservation of the environment], and were already recognized at a time when all we had to do with the world was to live in it together." Korsgaard (1993), pp. 50-51.

<sup>73</sup> T.M. Scanlon voices a similar view when he writes that the domain of "what we owe to each other...comprises a distinct subject matter, unified by a single manner of reasoning and by a common motivational basis. By contrast, it is not clear that morality in the broader sense is a single subject that has a similar unity." T.M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: The Belknap Press of Harvard University Press, 1998), p. 7.

The second component of Intersubjectivism is the heart of the view and is undoubtedly more controversial. Some readers may even take it to be a non-starter. How could our activities of granting one another practical authority actually make it the case that we *have* that authority? This scenario might seem to conjure normativity out of thin air, presupposing a more interactive but no less implausible variant of “thinking makes it so.” Let me here offer a few thoughts to persuade a skeptical reader not to hop off the train before it even leaves the station.

You might reasonably wonder whether thinking *ever* makes it so. Here’s a rough sketch of what we might call the *standard model*: on the one hand, there’s the world—“what is so”—and on the other hand, there are our beliefs about the world, with our chief epistemic norm directing us to apportion our beliefs to our evidence. On this model, we can at best hope that our thinking accurately reflects what is already, independently so. Our beliefs are responsive (if things go well), but never creative.

The standard model breaks down when we try to apply it to the social world. “Thinking makes it so” has significant creative potential when people are *thinking together*. Its creative potential becomes even greater when people use their intersubjective understanding to guide what they are *doing together*. I’d like to suggest that appealing to a practical, intersubjective version of “thinking makes it so” is, in fact, the only plausible strategy for explaining the existence and nature of complex social realities.

Consider what makes it the case that there are such things as friendships, laws, teams, nations, or economies. While the specifics differ from domain to domain, what creates and maintains all these social realities are people’s intersubjective attitudes and dispositions to act together. What makes it the case that two people are friends is a combination of facts about how they think, feel, and act toward one another over a span of

time, in light of each recognizing that the other has similar thoughts, feelings, and dispositions to act. Friendship is mind-dependent and intersubjective through and through. To take a very different example, what makes a \$10 bill worth something in a healthy economy—as a medium of exchange, a unit of account, and a store of economic value—is the shared belief that others also view it as worth something (namely, \$10) and guide their actions accordingly. These two examples suggest that we should set aside the worry that Intersubjectivism relies too heavily on the creative potential of “thinking makes it so.” Our intersubjective activities—our thinking and acting together—are responsible for creating and maintaining the entire social world.

Here an objector might pause. One could accept that our intersubjective activities have the creative power to determine “what is so” in the social world, yet resist the claim that they have the corresponding power to determine *what ought to be*. And Intersubjectivism’s second component posits something like this latter power, urging us to view moral normativity itself as intersubjective. Those inclined to resist this component of the view would most likely do so by objecting that intersubjective activities cannot account for the three features of moral normativity discussed above: its overriding importance, non-contingency, and authority.

I defend Intersubjectivism by making the case that it can account for these features, or near enough variations of them. The view treats moral claims as apt for normativity in virtue of the nature of the relationships in which we issue and respond to those claims. More specifically, it characterizes our moral thought and activity in terms of a trio of interdefinable concepts: being a person, being subject to the drive for respect, and being practically committed to participating in relationships of mutual accountability. In explicating these three concepts, Intersubjectivism aims to produce a conception of our moral lives that

comports with a plausible naturalistic conception while also being one that we can endorse practically. The possibility of this practical endorsement rests on how well the conception illuminates and coheres with our most central first-order moral beliefs. For instance, Intersubjectivism focuses on how the drive for respect structures our moral interactions and, in fact, some of our most central first-order moral beliefs are grounded in the idea that persons are obligated to act in ways that express respect for others. If Intersubjectivism can garner our practical endorsement, it will relieve the tension between our naturalistic and practical conceptions and thereby solve the reconciliation problem.

Before we proceed, it may be helpful to make a clarificatory point. It would be misleading to say that Intersubjectivism aims to derive moral normativity from a set of descriptive facts about our intersubjective activities. The prospects seem dim for that kind of project. Instead, Intersubjectivism aims to provide a psychologically rich characterization of our moral lives—as essentially intersubjectively connected with the moral lives of other persons—that also gives us reasonable confidence that our everyday moral interactions rest on a solid normative foundation. This foundation is intersubjective, but is nevertheless made solid by the fact that our very nature as persons disposes us to grant one another the practical authority that makes our moral interactions possible. To put the point a bit differently, we are not trying to show that normativity can be produced from something non-normative. Rather, we are trying to show how our everyday experience of moral normativity—our practical experience of moral claims exerting a normative grip on us—is intersubjectively supported by our shared human nature.

I build the foundation for Intersubjectivism in the next two sections. In §3, I argue that our drive for respect provides the fundamental psychological basis of our moral interactions. In §4, I defend an account of personhood that, if successful, will enable us to

defuse the worry that our intersubjective activities are too contingent and fallible to support moral normativity.

### 3. Living in the Minds of Others

#### 3.1 The Drive for Recognition

Let's begin by considering a fact that is puzzling, yet obvious upon reflection: a significant portion of our daily mental lives involves thinking about what other people think about us. For instance, we may want others to view us as clever, competent, considerate, or cute. The pervasiveness of these thoughts can seem puzzling in light of the fact that they often lack a straightforward instrumental justification, of the kind they would have if one were trying to impress others to gain some benefit from them. The pervasiveness of these thoughts is puzzling, in other words, because it's not immediately clear what could be valuable about appearing in another's mind in some particular way. Our persistent concern for others' opinions of us can therefore seem like nothing more than an unhealthy preoccupation. We would be more rational and autonomous—this line of thinking goes—not to mention less prone to anxiety and painful self-doubt, if we could simply free ourselves of the persistent concern for how we appear in the eyes of others. This set of ideas animates a central piece of mothers' wisdom: what matters is not that others hold good opinions of you, but that you actually *are* good and have an accurate sense of your worth.

These ideas also animated a number of Modern thinkers, who traced many of mankind's troubles back to what they variously referred to as *pride*, *vanity*, or *amour-propre*, a form of “self-love” that manifests as a psychological need to be thought well of by others.<sup>74</sup>

---

<sup>74</sup> I have in mind here thinkers such as Rousseau, Pascal, Montaigne, La Placette, Adam Smith, and Jacques Abbadie. For an excellent discussion of the role of the drive in 17<sup>th</sup> and 18<sup>th</sup> century thought, see A.O. Lovejoy, *Reflections on Human Nature* (Baltimore: The Johns Hopkins Press, 1961). Several

As their terms for it suggest, Modern thinkers conceived of our concern for others' regard chiefly as a vice.<sup>75</sup> But they took it to be a vice of a very special kind, for they identified it as the passion that most clearly distinguishes us from the other animals. For instance, Rousseau claims that being subject to *amour-propre* is what separates "sociable man" from "savage man." He writes,

[Sociable men are those who] count how they are looked upon by the rest of the universe [i.e., other people] for something, who can be happy and satisfied with themselves on the testimony of others rather than on their own. This, indeed, is the genuine cause of all [their] differences: the savage lives within himself; sociable man, always outside himself, is capable of living only in the opinion of others and, so to speak, derives the sentiment of his own existence solely from their judgment.<sup>76</sup>

Pascal joins Rousseau in portraying the concern to appear in others' thoughts as a universal and defining characteristic of humans. He writes,

We are not content with the life we have in ourselves and with our own existence; we wish to live an imaginary life in the thought of others, and we consequently force ourselves to appear.<sup>77</sup>

---

later thinkers—most notably Fichte, Hegel, Nietzsche, Sartre, and Husserl—also focus on our concern to appear in certain ways in the eyes of others, though their approaches differ in various ways from those of their predecessors. And, of course, many earlier thinkers discuss this trait in terms of people's concern to win *glory*, *reputation*, or *honor*, understood as excellent standing in the eyes of others. Aristotle, for instance, classifies the life of honor as one of the three main types of life. See Aristotle, *Nicomachean Ethics*, trans. Christopher Rowe and comm. Sarah Broadie (New York: Oxford University Press, 2002), I.5, pp. 97-98.

<sup>75</sup> Rousseau displayed much ambivalence about the effects of *amour-propre*, but his writings contain some of its most pessimistic treatments. For instance, he writes that the emergence of *amour-propre* "was the first step at once toward inequality and vice: from these first preferences arose vanity and contempt on the one hand, shame and envy on the other; and the fermentation caused by these new leavens eventually produced compounds fatal to happiness and innocence." Rousseau, *Discourse on the Origin and Foundations of Inequality among Men, or Second Discourse*, in Victor Gourevitch (ed.), *Rousseau: The Discourses and Other Early Political Writings* (New York: Cambridge University Press, 1997), p. 166.

<sup>76</sup> *Ibid.*, p.187. Although his rhetoric in the *Second Discourse* sometimes suggests otherwise, Rousseau does not hold that "savage man" is in fact a "man" at all, in the normal sense of the term. The primary aim of his hypothetical reconstruction of humanity's development is to show that the emergence of *amour-propre* is what gives rise to beings that are recognizably human.

<sup>77</sup> Blaise Pascal, *Pensées*, #806/147. Quoted in Lovejoy (1961), p. 229.

Both thinkers distinguish between living within oneself—characterized as a virtue of which humans are not fully capable—and living outside of oneself—characterized as a vice susceptibility to which is definitive of being human at all.

Why might it be a vice to seek “an imaginary life” in the thought of others? The piece of mothers’ wisdom above offers us a hint: how one appears to others does not matter, in and of itself. So, at best, the energy one expends seeking to appear to others interferes with the energy one could expend pursuing things that actually matter. At worst, seeking to appear to others could corrupt one’s ability to pursue or accurately judge what is valuable at all. For instance, it could motivate one to act in accordance with others’ evaluative standards in order to win their esteem, or it could lead one to value mere appearance over reality.<sup>78</sup> This last worry may seem especially germane to contemporary readers, who encounter a social world with ever more incentives to sacrifice having some valuable experience in order to merely appear to others as though they are having that valuable experience.<sup>79</sup>

Rousseau and Pascal also agree in emphasizing that we cannot be content as long as our concern to appear to others is frustrated. Here is Pascal again:

---

<sup>78</sup> Rousseau holds that proper moral education is necessary to guard against the first possibility. For instance, he writes, “Emile...values nothing according to the price set by opinion; thus, although he likes to please others, he will care little about being esteemed by them...He will not precisely say to himself, ‘I rejoice because they approve of me,’ but rather, ‘I rejoice because they approve of what I have done that is good.’” Rousseau, *Emile, or on Education*, trans. Allan Bloom (New York: Basic Books, 1979), pp. 338-39. Regarding the second possibility, Rousseau claims that with the genesis of *amour-propre*, “To be and to appear to be became two entirely different things, and from this distinction arose ostentatious display, deceitful cunning, and all the vices that follow in their wake.” Rousseau (1997), p. 170.

<sup>79</sup> An example I have observed: a man laying on a beach does not relax while watching the waves roll in, but rather spends ten minutes trying to get the best picture of himself as he pretends to relax while watching the waves roll in. Upon uploading the picture to social media, he continues to ignore the scene in favor of checking his phone for signs of others’ approval.

The quest for glory is the quality that is most ineffaceable from the heart of man... However much of health and of essential comforts he may have, he is not satisfied unless he have a place in men's esteem.<sup>80</sup>

Pascal suggests that just as we have physical needs for things like food and shelter, we have a kind of *psychological need* to know that others think well of us. If we accept this claim and combine it with the observations that our concern for others' regard is persistent, pervasive, and universal among normal humans, it becomes clear that the relevant psychological mechanism is better classified as a *drive* than as a mere desire.

Accordingly, let's call the *drive for recognition* the psychological force that motivates us to seek the positive regard of other people. More specifically, we are interested in the following form of the drive:

**Drive for Recognition:**

The psychological force that leads a person, *A*, to want another person, *B*, to take *A* as the object of some attitude, *x*, because (i) *A* believes that (for *B*) *x* expresses positive regard, and (ii) *A* values *B*'s positive regard (at least partly) independently of any instrumental benefit *A* could gain from receiving it.

We can think of the drive for recognition as a species of the more general *drive for inclusion* that we share with other social animals. What makes these animals social is their drive to seek interactions that signal inclusion, such as through proximity, attention, communication, affection, play, displays of dominance or deference, and so on. The drive for recognition leverages humans' more sophisticated cognitive capacities to make the kinds of social interactions we seek psychologically distinctive. We can better understand what is special about the drive for recognition by thinking through how these sophisticated capacities enable its operation.

Five capacities are particularly important. First, we have the conceptual capacity to classify particular things by understanding how their qualities distinguish them as things of a

---

<sup>80</sup> Pascal, *Pensées*, #470/404. Quoted in Lovejoy (1961), p. 132.

certain type. For instance, we can identify some object as a spear by noting how its aerodynamic shape, tapered tip, and manageable heft make it a useful tool for hunting. Second, we have the evaluative capacity to assess these particular things according to standards. Exercising these two capacities in tandem enables us to apply concepts like *good* or *bad* to particular things.<sup>81</sup> The content of our evaluative standards is often tied to the very qualities we use in classifying their objects. This is most obviously the case with human artifacts, the qualities of which we have designed with specific ends in mind. What determines whether some spear is good is how well it can achieve the purpose of a spear, such as by having a tip sharp enough to mortally wound, a heft light enough to heave yet heavy enough to fly true, and durability sufficient to last many hunts.<sup>82</sup>

Third, our capacity for self-consciousness enables us to exercise these first two capacities—for classification and evaluation—in thinking about aspects of *ourselves*. In

---

<sup>81</sup> Other animals must employ simpler versions of these capacities for classification and evaluation. My dog certainly distinguishes things that are food from things that are not food. Moreover, when faced with a choice between two types of food (say, on two separate plates of scraps) she clearly acts on the basis of preferences for one over another, which suggests that she employs some kinds of evaluation. A plausible account of what distinguishes human versions of these capacities would appeal, at minimum, to the conceptual sophistication of our classifications and to our explicit use of evaluative concepts. And a plausible account of how these capacities have developed would appeal, at minimum, to how our dispositions to cooperate have made possible cumulative cultural evolution and various kinds of social institutions. A plausible account of *these* phenomena would appeal, at minimum, to the three other capacities underlying the drive for recognition that I go on to discuss. For illuminating discussion of several of these issues, see Michael Tomasello's *The Cultural Origins of Human Cognition* (Cambridge, MA: Harvard University Press, 1999) and *Why We Cooperate* (Cambridge, MA: The MIT Press, 2009).

<sup>82</sup> The characterization in this paragraph shares much in common with Aristotelian views of what individuates particular objects and what standards govern our assessments of those objects. Christine M. Korsgaard defends such a view when she writes,

Teleological organization is what unifies what would otherwise be a *mere heap* into a particular object of a particular kind...To know an object, that is, to *understand* it, is to see not only what it does and what it is made of, but also *how* the arrangement of the parts enables it to do whatever it does...At the same time, it is the teleological organization or form of the object that supports normative judgments about it.

(Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (New York: Oxford University Press, 2009), p. 28.) It is easier to accept this view about human artifacts than about some other things, such as persons themselves. The point I am trying to establish in the present discussion is simply that our capacities for classification and evaluation—however they are best characterized—are necessary for the drive for recognition.

particular, we can form conceptions of ourselves that reflect the characteristics we care about, such as the social roles we play and the traits of which we are proud or ashamed. For instance, it may be part of a person's self-conception that she is a good hunter, consistently able to down a wild boar with the throw of a single spear. It is crucial to notice that two levels of evaluation are at work in constructing a self-conception: a person can judge whether she is a good hunter, as well as whether it is good for her to be a good hunter. After all, it might be the case that she believes she is a good hunter but that she attaches no value to that characteristic, since she considers farming and trade to be the only noble ways of procuring food. So a person's self-conception includes characteristics assessed according to standards particular to each characteristic—a talented hunter, a middling farmer, a hapless trader—as well as attitudes about which characteristics are worth having. The form of the drive for recognition we are studying seeks positive regard, and so concerns only those characteristics a person thinks it good to have.<sup>83</sup>

The dynamics through which we develop and modify our self-conceptions are intimately bound up with our exercise of a fourth capacity: we are aware of other people as subjects just like us, who can take *us* as the objects of their own classification and

---

<sup>83</sup> Of course, an alternative form of the drive for recognition could seek another's *disapproval* for some characteristic of which one is ashamed, as one may simply want another to confirm one's low opinion of oneself. A trickier case for the form of the drive I am discussing involves a person seeking attitudes that are not usually considered forms of positive regard, but that others express precisely because they value some characteristic the person possesses. For instance, a painter may want her rivals to feel envy or hatred for her when they see her impressive new body of work, precisely because they recognize how excellent a painter she is. Similarly, an athlete may want her opponents to experience terror at the prospect of facing her in competition. The key to interpreting these cases lies in understanding that the others' apparently negative attitudes (such as envy, hatred, and terror) are second-order reactions to their attitudes of positive regard (such as awe, appreciation, or admiration). The tennis player would not fear stepping on the court against her opponent if she did not think highly of her opponent's skills.

evaluation.<sup>84</sup> Exercising this capacity makes possible various forms of intersubjective awareness, wherein *A* recognizes *B* as a subject who can, in turn, recognize *A* as a subject who can recognize *B* as a subject, and so on. The proud hunter can imaginatively project herself into the minds of others to determine whether they share her assessment of her spear-throwing prowess, as well as whether they think highly of having that prowess. The fifth capacity underlying the drive is perhaps the least obvious, but will be crucial in the arguments to come. In seeking recognition from another for some characteristic, one implicitly grants the other a kind of authority. Stated differently, valuing another's positive regard for some characteristic presupposes that the other is a competent judge both of that characteristic and of the value of possessing it.

The operation of these five capacities—for classification, evaluation, self-consciousness, intersubjectivity, and the granting of authority—distinguishes the drive for recognition from the more general drive for inclusion. I next want to consider two questions that arise from Modern thinkers' treatment of the drive. First, are they right to conceive of the drive for recognition chiefly as a vice? Second, are they right to portray one's being subject to the drive as essential to one's being a person at all? In the next section, I argue that focusing on the drive's vice-like qualities conceals a deeper, more interesting truth about the role it plays in making us moral beings. Thus, I argue that we should answer the first question with a qualified "no" and the second with a resounding "yes."

---

<sup>84</sup> C.H. Cooley articulates the notion of the "looking-glass self" to explain how our sensitivity to others' judgments about us leads us to construct our self-understandings in concert with how (we think) they see us. See Cooley, *Human Nature and the Social Order* (New York: Scribner's, 1902).

### 3.2 Two Forms of Positive Regard

I believe that Modern thinkers viewed the drive for recognition chiefly as a vice because they tended to focus on manifestations of the drive that seek a particular type of positive regard—namely, esteem. Let's say that the *drive for esteem* seeks regard for valued characteristics that a person may possess to a greater or lesser degree. Esteem is socially comparative. It picks out and evaluates those characteristics that distinguish a person as the *particular* person she is, or at least as the particular person another takes her to be, and thereby expresses a conception of a person's relative standing among others. I believe that Rousseau, Pascal, and our mothers are right to warn us about the dangers of the drive for esteem. Allowing this drive free reign *can* encourage corruption, superficiality, duplicity, ruthless competition, and widespread dissatisfaction. These last two phenomena are mutually reinforcing. Even those fortunate enough to win others' esteem can never be fully secure in their standing, since they are always vulnerable to losing their esteemed trait, being supplanted by others who possess the trait to a greater degree, or having the evaluative standards that govern others' esteem shift entirely. Beauty fades, new beauties come along, and the standards of beauty themselves vary across time and place.

Nevertheless, the drive for esteem also has some redeeming features. Here I mention three. First, the drive can play an important epistemic role in enabling us to corroborate our self-conceptions by appealing to others' views of us. This is the feature Aristotle has mind when he observes that

...People seem to pursue honor in order to be convinced that they themselves are good: at any rate they seek to be honored by people of discernment, and among those who know them, and to be honored for excellence.<sup>85</sup>

---

<sup>85</sup> Aristotle (2002), I.5, p. 98.

On this view, rather than leading us to abandon the pursuit of what is good in favor of merely appearing good, the drive for esteem can aid our attempts to be good. Seeking esteem from those who are well placed to judge whether we actually possess estimable qualities allows us to use their expressed (or withheld) esteem as evidence that informs our self-assessments. Second, rather than corrupting our evaluative standards, the drive for esteem can supplement our motivations to act well and to pursue what is good (insofar as we seek esteem from people who hold good values). Or, to approach the point from the other direction, the prospect that others may think less well of us if we act poorly can motivate us to act well. The composer may devote more time to perfecting her sonata, for instance, if she knows her audience will include people whom she wants to impress and whose derision she fears. More broadly, seeking esteem from people who value good things that we do not yet value may inspire us to incorporate their values and seek to live up to them. Third, it is reasonable for a person to care that the particular conception under which she values herself aligns with the conceptions under which others value her. A person can suffer from a kind of social alienation if, say, she cares most about being witty and athletic but others esteem her only for her wealth, or fail to esteem her at all.

These three salutary features of the drive for esteem suggest that we should not classify it straightforwardly as a vice, even though it often manifests in problematic ways. If this is right, then we can offer a partial answer to the puzzle with which we began the previous section—namely, how to make sense of the incessant activity of the drive for recognition in light of the idea that what the drive seeks is worthless. The partial answer maintains that what the drive seeks is not always worthless. A full answer would improve on the partial answer by addressing two of its shortcomings. First, it could still be true that the drive for esteem exerts a negative influence that outweighs its positive contributions, in

which case we would still have reason to be ambivalent about the role the drive for recognition plays in our psychological lives. Second, the partial answer succeeds, insofar as it does, only by denigrating the piece of mothers' wisdom, which would disappoint our mothers. Thankfully, we can address both shortcomings at once. For we have at our disposal a second piece of mothers' wisdom that we can endorse wholeheartedly and that illuminates how the drive for recognition enables us to function as persons.

Consider the mothers' imperative: "Stand up for yourself." Even if we should avoid worrying about whether others think we are cool or beautiful, we should not let others walk all over us or treat us as though we matter less than they do. The mothers' imperative instructs us to seek a second form of positive regard: *respect*. More specifically, it tells us to demand from others the respect that is owed to us simply as persons. In standing up for oneself, one *asserts one's standing* as a person. We might say that this standing is one of equal footing with other persons, and that it enables one to look others eye-to-eye. The etymology here is suggestive. The Latin *respectus* means "the act of looking back at one," and the 14<sup>th</sup> century French *respecter* means "to treat with deferential regard."<sup>86</sup> The very act of looking at a person in the right way, of *recognizing her as a person*, leads other persons to grant her standing in their thought and action. The mothers' imperative comes into play because it may be one's very assertion of one's standing—the act of standing up for oneself—that forces the other to look at and treat one in the right way, i.e., as a person.

Respect for persons is a distinctively moral attitude. Yet we can bring out two of its most important features by examining the character of other, non-moral attitudes that are also commonly thought of as forms of respect.<sup>87</sup> For instance, we can respect the power of a

---

<sup>86</sup> See <http://www.etymonline.com/index.php?term=respect>.

<sup>87</sup> Several of the following instances of non-moral respect might also call for one's moral respect, depending on how the details are spelled out.

raging river by deciding not to risk wading across it, and we can respect the value of the river itself by not dumping chemicals into it. We can respect the Dalai Lama by referring to him as “His Holiness” and by not sitting in a seat taller than the one in which he sits. We can respect people’s legal rights by not trampling them and by ensuring that unfair obstacles do not impede their exercising those rights.<sup>88</sup> Despite this diversity of use, in nearly every case we express our respect for an object not simply through our attitudes, but also through our *actions*. This intimate connection to action distinguishes respect from esteem.<sup>89</sup> My admiration for Stephen Curry’s jump shot needn’t lead me to *do* anything. Of course, I could express my admiration by cheering while watching him play or by buying a replica of his jersey to wear around town. If Curry were seeking esteem from me, these are the kinds of behaviors he would look for as indications that I do, in fact, hold his shooting abilities in high regard. While these actions are outward signs of my esteem, they are neither essential to my esteem, nor are they necessarily directed toward the object of my esteem. That is to say, I could esteem Curry without ever showing it, and even if I did show it, I needn’t show it to Curry himself.

Things are quite different in the case of respect. There is essentially no gap between respecting an object and expressing respect through how one behaves toward that object. For instance, I express my respect for a person’s property by, at minimum, treating the fact

---

<sup>88</sup> On this last point, I have in mind things like laws intended to discourage members of minority groups from exercising their right to vote, or grinding poverty functioning as an obstacle to people’s exercising their rights to free movement, private property, and much else.

<sup>89</sup> Stephen Darwall and Frederick Neuhouser have also emphasized how the essentially practical character of respect distinguishes it from esteem (roughly aligning with the distinction between what Darwall refers to as “recognition respect” and “appraisal respect”). See Darwall, “Two Kinds of Respect,” *Ethics* 88 (1977), pp. 36-49, and Neuhouser, *Rousseau’s Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition* (New York: Oxford University Press, 2008), pp. 61-63.

that it is hers as a side constraint on how I act toward it.<sup>90</sup> I do not steal or damage her property, and I do not use it without her permission. This asymmetric relevance of action reflects a fundamental difference in the types of judgments that underlie esteem and respect, respectively. While esteem is an evaluative form of thought, respect is a norm-directed form of thought. It aims to guide our actions in response to purported reasons. Respect organizes and expresses our practical relationships to its objects by providing principles intended to direct our actions with regard to those objects. Accordingly, let's call this the *norm-directed* aspect of respect.

To be clear: in characterizing the norm-directed aspect of respect, I am aiming to classify a form of thought. I am not taking a stand on whether the norms to which respect directs us are themselves normative. That determination—that an object is *deserving* of respect—could only be made on a case-by-case basis by engaging in first-order moral reasoning. My point here is that in expressing respect, one will necessarily take the object to possess features that (normatively) call for one to respect it.

When we shift our attention from the *subjects* of esteem and respect to their respective *objects*, we notice that the asymmetric relevance of action is often reversed. Many of the characteristics we esteem are achieved only through action that distinguishes people from one another. Quite often, this action demonstrates uncommon talent and perseverance, which helps to explain why esteemed characteristics are embodied to varying degrees and distributed unevenly. By contrast, the objects that seem to call for our respect do so simply in virtue of being the kinds of objects they are. Once these objects meet certain criterial thresholds, we take them to be deserving of our respect. We respect the raging river because it is dangerous, and we respond to this same feature in respecting the active lava

---

<sup>90</sup> The concept of rights as side constraints comes from Robert Nozick. See Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 30-33.

flow or the unstable slope at risk of avalanche. We respect people's legal rights in virtue of their possessing the set of characteristics necessary for that status, such as being over a certain age, being of sound mind, and being a legal citizen. We also respect people in virtue of their inhabiting certain general roles, such as those of parent or child, or that of a client to whom one has assumed a fiduciary duty.<sup>91</sup>

These examples suggest that the objects of our respect purport to call for our practical consideration simply in virtue of their being the kinds of objects they are, not in virtue of having done anything to distinguish themselves from other objects of their kind. We might capture the difference, imperfectly but pithily, like this: we esteem characteristics that determine comparative standing and respect characteristics that determine general standing. In respecting an object, one grants it the standing to make certain claims on one's attitudes and actions simply because it is an object of the type it is.<sup>92</sup> Accordingly, let's call this the *general standing* aspect of respect.

Respect for persons exhibits the norm-directed and general standing aspects just discussed, but it differs from other forms of respect in being distinctively moral. To respect someone as a person is to grant her the moral standing to make claims on the attitudes and actions of other persons, and to grant her this standing just in virtue of the fact that she is a person. While esteem attaches to characteristics that distinguish persons from one another, respect is deemed warranted on the basis of characteristics that all persons share. Different moral theories advance different characteristics to fill the latter role. Some focus on our

---

<sup>91</sup> Of course, the fact that one is *in* such a role often reflects past actions, decisions, and commitments. The point is simply that one's being in that role is what (we think) makes one deserving of the appropriate form of respect, and that one needn't do anything that counts as excelling in that role in order to (we think) deserve respect.

<sup>92</sup> The notion of an object's "making a claim" must be understood broadly. A person will interpret the danger posed by the rushing river as making a claim on her attitudes and behavior insofar as she is prudentially rational, since its being dangerous gives her prudential reason not to risk wading across it, or to take extra care when doing so.

rationality, others on our autonomy, others on our capacity to act morally, yet others on each person's ability to pursue a life plan that reflects his or her conception of the good.<sup>93</sup>

Many of these theories are developments of or reactions to Kant's view, which manages to incorporate versions of all of these characteristics by arguing for their essential interrelation:

...[A] human being regarded as a *person*, that is, as the subject of a morally practical reason...is not to be valued as a means to the ends of others or even to his own ends, but as an end in himself, that is, he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world. He can measure himself with every other being of this kind and value himself on a footing of equality with them.<sup>94</sup>

Kant holds that persons are owed respect in virtue of their dignity as ends in themselves.

The dignity of persons depends on their autonomy, which Kant understands as their ability to confer value on their ends by determining their wills according to universal laws. Since persons are the ultimate sources of value—as self-legislating ends in themselves—we cannot universalize any maxim that fails to treat persons with respect, i.e., by treating them as a mere means to our ends.<sup>95</sup>

My main interest here is not to grapple with the complexities of Kant's substantive view of why persons are owed respect. Rather, I want to draw out the broad ways in which

---

<sup>93</sup> For a statement of this last view, see John Rawls, *A Theory of Justice, Revised Edition* (Cambridge, MA: The Belknap Press of Harvard University Press, 1999), §63, p. 358.

<sup>94</sup> Immanuel Kant, *The Metaphysics of Morals*, trans. Mary Gregor (New York: Cambridge University Press, 1996), 6: 434-35, p. 186.

<sup>95</sup> This is how I interpret Kant's argument that

...[M]orality is the condition under which alone a rational being can be an end in itself, since only through this is it possible to be a lawgiving member in the kingdom of ends. Hence morality, and humanity insofar as it is capable of morality, is that which alone has dignity...For, nothing can have a worth other than that which the law determines for it. But the lawgiving itself, which determines all worth, must for that very reason have a dignity, that is, an unconditional, incomparable worth; and the word *respect* alone provides a becoming expression for the estimate of it that a rational being must give. *Autonomy* is therefore the ground of the dignity of human nature and of every rational nature.

Kant, *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor (New York: Cambridge University Press, 1998), 4: 435-36, pp. 42-43.

his view coheres with my descriptive account of what is involved in the activities of granting or claiming respect. The fact that there is such broad coherence among the two is a testament to Kant's success in capturing central features of our moral experience, as well as to his lasting influence on our everyday moral thinking. Notice how Kant's view incorporates the norm-directed and general standing aspects of respect. A person's dignity allows him to "exact respect for himself" from other persons by placing limits on how others may treat him, i.e., never merely as a means. And respect for persons is socially comparative only in the minimal sense that it treats all who qualify for it as *equally* qualified. Every person is equally a person—on a "footing of equality," as Kant says—and so has equal standing to demand the respect owed to persons. Also notice how Kant identifies deep connections between being a person, being a moral subject, and being disposed to exact respect for one's standing as a person.

In the following section, I argue that our everyday practices of holding one another morally accountable presuppose some version of this connection, even if many people have attempted to make sense of these practices under alternative conceptions. I defend an account that focuses on the conceptual and functional ties between treating one as a person and granting one the standing to demand respect. In so doing, I advance a conception of personhood according to which it is essential to being a person that one has the capacity to participate in relationships of mutual accountability with other persons. And participating in these relationships, I argue, is governed by the drive for respect. On my view, persons are agents whose drive for respect leads them to engage in moral relationships.

The success of my arguments would help provide what I above referred to as a "full answer" to the question about the worth of the drive for recognition. While the drive for esteem can manifest in damaging ways, the drive for respect enables us to participate in

moral relationships and thereby makes us persons.<sup>96</sup> In playing this role, the drive for respect undermines the dichotomy between reality and mere appearance that gave rise to worries about the value of the more general drive for recognition. It also addresses the initial worry that our intersubjective activities could not account for morality's overwhelming importance. Our mutual granting of the standing to make moral claims on one another—which largely consists in figuring in others' thoughts in various ways—is what makes moral relationships possible. If these claims are true, then Modern thinkers were right to identify the drive among our most essential characteristics as persons.<sup>97</sup>

#### **4. How the Drive for Respect Makes Us Persons**

##### **4.1 Persons as Second-Persons**

Suppose that you are skeptical of my claim that moral relationships depend on our being subject to the drive for respect. One source of skepticism might be your first-order moral view. My claim above is a descriptive one about how our psychologies lead us to relate to and interact with one another. But it is also intimately bound up with how we conceive of the normative basis of those relations and interactions. Our first-order moral beliefs guide how we interact morally, and so are necessarily reflected in our descriptive conception of what those interactions are about. Suppose, then, that you endorse a first-order moral view according to which morally right actions are those that (aim to) bring about good outcomes,

---

<sup>96</sup> Compare this claim to Rousseau's view:

So long as [one's] sensibility remains limited to his own individual being, there is nothing moral in his actions. It is only when his sensibility begins to extend outside himself that it takes on, first, the sentiments and, then, the ideas of good and evil which truly constitute him as a man and an integral part of his species.

Rousseau, *Emile, or On Education*, (trans.) Allan Bloom (New York: Basic Books, 1979), pp. 219-20.

<sup>97</sup> A fully satisfying account would also explain how the drive for respect emerges from the more basic drive for esteem, both developmentally as a human individual becomes a person and over the course of human history as individuals begin to employ concepts of respect within relationships of mutual accountability.

rather than those that (aim to) relate us to other persons in particular ways—specifically, those that express respect for others’ standing or that demand they grant us the reciprocal respect.

I’d now like to show that your everyday moral experience speaks against the former conception and in favor of the latter one. This divergence wouldn’t prove that your first-order moral conception is false, of course. That is a normative matter. But it would reveal a troubling schism between the moral view you endorse and the purportedly moral considerations that actually guide your attitudes and actions. Moreover, for the purposes of assessing Intersubjectivism, it is important to capture what is actually going on with our minds, since the view argues that how our minds work (together) bears on the nature of moral normativity itself.

Consider the following scenario:

*Impersonal Pizza:* You are hurrying down the sidewalk carrying a freshly baked pizza, eager to get home and eat a slice. You fail to notice a raised seam in the sidewalk ahead of you. Stubbing your toe on the seam, you lurch forward and flail your arms to break your fall, launching the pizza box into a nearby puddle. Your pizza is ruined.

This is a bad outcome. You were going to enjoy eating that pizza. You will reasonably feel disappointed, and you will likely regret not having been more attentive, since you might have saved yourself from stumbling. Now consider a second scenario:

*Personal Pizza:* You are hurrying down the sidewalk carrying a freshly baked pizza, eager to get home and eat a slice. A passerby slaps the pizza box out of your hands, knocking it into a nearby puddle. Your pizza is ruined. The passerby laughs and continues on her way.

This is an equally bad outcome. In both cases, your pizza is ruined. But consider how your reaction in the second scenario would differ from your reaction in the first. In Personal Pizza, the bad outcome is the result of the passerby wronging you, rather than the result of a mere accident. You will again feel disappointed that you will not be able to enjoy the pizza.

But now your disappointment will likely be paired with a second attitude. You will likely *blame* the passerby for the disrespect she has shown you, and likely express your blame through *resentment*. You might also demand an explanation, an apology, or restitution. It is important to see that your resentment is not merely a response to the belief that the passerby acted wrongly by intentionally causing a bad outcome, or even that the passerby acted wrongly by intentionally harming another person, as though what she did were some fact independent of how she relates to you. Your resentment is a response to the belief that the passerby has *wronged you*, failing to exhibit the respect she owes you as a fellow person.

I submit that in Personal Pizza, your resentment of the passerby's disrespect is likely to loom significantly larger in your mind than is your disappointment at suffering the loss of your pizza. Rousseau was sensitive to this phenomenon:

As soon as men had begun to appreciate one another and the idea of consideration had taken shape in their mind, everyone claimed a right to it, and one could not deprive anyone of it with impunity... [A]ny intentional wrong became an affront because, together with the harm resulting from the injury, the offended party saw in it contempt for his person, often more unbearable than the harm itself.<sup>98</sup>

Following Rousseau, I contend that the drive for respect shapes our shared understanding of what morality requires, thereby mediating our attitudes toward one another. What matters to us in everyday life is not just that people cause bad outcomes by failing to live up to morality's requirements. Our chief moral concern is that people disrespect others by failing to treat them with the consideration we think they are owed simply as persons.<sup>99</sup>

---

<sup>98</sup> Rousseau (1997), p. 166.

<sup>99</sup> Adam Smith argues for a similar view: "[What most] enrages us against the man who injures or insults us is the little account which he seems to make of *us*...that absurd self-love, by which he seems to imagine, that other people may be sacrificed at any time, to his conveniency." Adam Smith, *Lectures on Jurisprudence*, ed. R.L. Meek, D.D. Raphael, and P.G. Stein (Indianapolis: Liberty Classics, 1982), p. 96. Quoted in Stephen Darwall, *The Second-Person Standpoint: Morality, Respect, and Accountability* (Cambridge, MA: Harvard University Press, 2006), p. 84.

This conception of our moral concern is *second-personal* in the sense discussed by Stephen Darwall. He calls “*the second-person standpoint* the perspective you and I take up when we make and acknowledge claims on another’s conduct and will,” and he claims that “[o]ur dignity as persons includes...an irreducibly second-personal authority to demand respect for this very authority and for the requirements which it gives us standing to demand compliance.”<sup>100</sup> The standpoint is second-personal because it casts each person as a “you” to whom certain forms of treatment can be owed in virtue of each person’s authority, and not simply as a “he” or a “she” whom there is (third-personal) reason to treat in some ways and not in others.

On my view, persons are *essentially* second-persons. Our status as persons depends upon our having the capacity to participate in moral relationships with others by holding ourselves and others accountable to the standards of those relationships. These standards operate on the mutually granted authority of each person to make claims on others. Persons thus understand themselves as sources of valid claims on other persons,<sup>101</sup> where this self-understanding includes the disposition to exercise these claims in the appropriate situations, and they view other persons as having the same standing with regard to them. This conception views persons as essentially related to one another through the mutual granting of moral standing. Moral engagement and personhood are thus inter-defined in terms of this distinctive relationship of mutual accountability.<sup>102</sup>

---

<sup>100</sup> Darwall (2006), pp. 3, 14.

<sup>101</sup> I borrow part of this phrase from Rawls, who characterizes persons as “self-originating sources of valid claims.” I leave out “self-originating,” however, because I believe that a person’s being a source of valid claims on others is ultimately an intersubjective phenomenon. See Rawls, “Kantian Constructivism in Moral Theory,” *Journal of Philosophy* 77 (1980), pp. 515-72.

<sup>102</sup> My conception of personhood has been significantly influenced by Darwall’s similar conception, though his is straightforwardly normative. Darwall writes,

...[T]o be a person just is to have the competence and standing to address demands as persons to other persons, and to be addressed by them, within a community of mutually

## 4.2 Accountability and Norm-directed Expectations

I will now make the case that our drive for respect is what enables us to participate in relationships of mutual accountability, and thus what makes us persons. Our practices of holding one another accountable rest on a foundation of what I call *norm-directed expectations of regard*. We expect certain kinds of regard—in some cases respect, in other cases forms of consideration or affection—from the people with whom we interact. The content of these expectations is governed by the standards of the various forms of relationship in which we participate. For instance, the sort of regard we expect from an acquaintance to whom we're telling a story in a (normal, non-coerced) conversation—say, to make eye contact regularly, to demonstrate some minimal degree of interest, and to not start telling their own story in the middle of our own—is governed by norms of conversation and etiquette which, when followed, express a form of respect.

These expectations of regard are norm-directed in the sense that we don't simply presume that others will, by and large, exhibit the relevant attitudes and behaviors; we also hold that they *should* exhibit these attitudes and behaviors. When they fail to do so in instances that constitute moral disrespect, our norm-directed expectations tend to be expressed as *demands* that others comply with the relevant norms, as well as that they apologize, show compunction, or in some other way make amends for their exhibited lack of regard. As I hope this sketch makes clear, norm-directed expectations are not necessarily normative, full stop. The fact that one holds a norm-directed expectation does not, on its own, generate reasons for others to satisfy that expectation. Think, for instance, of the

---

accountable equals..We therefore respect another as a person when we accord him this standing in our relations to him.”  
Darwall (2006), p. 126.

norms surrounding dueling in response to a challenge or insult. Rather, norm-directed expectations express individuals' judgments about the norms with which persons must comply in order to show appropriate respect for others.

Holding others and ourselves to norm-directed expectations is constitutive of our participation in relationships of mutual accountability because it is via these expectations that we hold others accountable. While my main focus is the basic moral relationship, the conversational etiquette example illustrates how norm-directed expectations govern a variety of everyday interpersonal relationships. Our more or less consistent, more or less mutual expression of the appropriate forms of regard is what structures and expresses the character of those relationships. For example, think of how the sorts of regard we expect from strangers on the street—say, not to steal from us or knock our pizzas into puddles, and to lend us needed help if they can do so at no serious cost to themselves—define and organize the nature of our interactions as members of a local community.

And think of how something analogous can be said of the very different sorts of regard we expect from close friends—say, to be excited at the prospect of spending time with us, to care about how our lives are going, to keep our confidences, and to want to comfort us when we are having a difficult time. The regular, mutual exhibition of these qualities is what (among other things) *makes it the case* that the two people are close friends. If one or both of them fall short of the relevant norm-directed expectations, their friendship is likely to be impaired. And if one or both fall short in some extreme way, they may no longer be close friends at all. We can see, then, that these norm-directed expectations set forth both

the ideal of the good close friend and the conditions under which someone counts as a close friend at all, i.e., even when one fulfills that role less than ideally.<sup>103</sup>

Something analogous is true of the moral case: our norm-directed expectations set forth the ideal of the good moral agent and presume, as a condition of their felicity, that their object possesses the capacities necessary for moral agency at all. It is in employing norm-directed expectations of regard that we hold others and ourselves morally accountable. We tend to blame those who treat us in ways that fall short of our norm-directed expectations and act gratefully toward those who display degrees of regard that go above and beyond our expectations. If someone failed to have and regularly act on the disposition to hold herself and others accountable in these ways, then, absent some good explanation, it would be difficult to see how she could count as a competent participant in moral relationships.

Of course, there are several contexts in which we do have such a “good explanation” for not holding another accountable for behavior we would normally deem disrespectful. I’ll here mention five types of case. First, and most straightforwardly, a person may determine that another’s behavior does not exhibit a lack of respect, despite initial appearances. For instance, we might excuse the passerby in Personal Pizza if we realized that she knocked our pizza to the ground by accident in the process of trying to break her own fall. Second, a person may decide that another’s disrespectful behavior is properly excused due to mitigating circumstances. For instance, we might discover that while the passerby did ruin our pizza intentionally, she has been under unusual amounts of stress lately, or has recently suffered some personal tragedy, such that her behavior is not an accurate reflection of her

---

<sup>103</sup> T.M. Scanlon has argued for essentially the same way of thinking about the norm-governed attitudes and dispositions that define personal relationships. For his account of friendship, see especially *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, MA: Belknap Press of Harvard University Press, 2008), pp. 131-36.

“real” or typical attitude toward us. Third, one might determine that even though another’s apparently disrespectful behavior *is* an accurate reflection of her “real” attitudes, she is not a proper object of one’s norm-directed expectations because she lacks the capacities to satisfy them. For example, one might discover that the passerby suffers from severe psychological deficiencies.

The final two types of case are more fraught, because they reflect a more systematic breakdown of moral relations. In the fourth type of case, a person may fail to demand respect because she knows that the other will in fact fail to express it, or because demanding respect will foreseeably have bad consequences. For instance, think of a slave considering whether to demand respect from a slave-trader. In the fifth type of case, a person may have internalized a demeaning self-conception, according to which she lacks the standing to make certain moral claims on those she considers her superiors. For instance, think of the restrictive effects of taking oneself to be a member of a lesser gender, ethnicity, class, or caste. In the fourth case, one is unable to fully participate in relationships of mutual accountability because others systematically deny one standing, while in the fifth, one is unable to participate because one does not take oneself to have that standing, and so does not hold others accountable for granting it. If persons are essentially second-persons, as I have been arguing, then these two cases suggest that especially poisonous social environments may prevent some of their members from realizing their full potential as persons.

#### **4.3 Strawson on Personhood and Reactive Attitudes**

My general approach to understanding our practices of holding one another morally accountable is indebted to the work of Darwall, T.M. Scanlon, and Peter Strawson, among

others.<sup>104</sup> It will be helpful to say a bit more here about Strawson's account, for two reasons. First, his emphasis on the essentially affective dimension of holding accountable is of a piece with my claim that our concern for respect issues from a *drive*. Second, he takes a similar perspective on (what I've referred to as) the inter-definability of personhood, the drive for respect, and participation in relationships of moral accountability.

Strawson famously argues that to hold another person accountable just is to be prone to form a reactive attitude toward her when you judge that she has failed to demonstrate some appropriate form of regard (or "goodwill," as Strawson calls it), or when you judge that she has demonstrated a form of regard over and above what is called for. Examples of reactive attitudes include resentment, indignation, gratitude, and, in the reflexive case, guilt. On Strawson's account, we hold persons accountable by reacting to their interpersonal attitudes with a range of distinctively affective interpersonal attitudes (and sometimes intrapersonal attitudes) of our own. As I've discussed above, to hold a passerby accountable for ruining your pizza is to be prone to resent her for having done so. Your resentment is *reactive* in the sense that it responds to your judgment about the quality of the other person's attitudes—in this case, her lack of respect for you and your right to transport your pizza unmolested.

Above, I stressed the significant difference between resenting someone for ruining your pizza and merely regretting the badness of the fact that your pizza has been ruined.

Consider a third scenario:

*Unbearable Pizza:* You are hiking down a wooded trail carrying day-old pizza, eager to get to your campsite and eat a slice. A bear emerges from the brush. She is also eager

---

<sup>104</sup> Darwall (2006). Peter Strawson, "Freedom and Resentment," *Proceedings of the British Academy* 48 (1962), pp. 1-25. Reprinted in Gary Watson (ed.), *Free Will* (New York: Oxford University Press, 2003), pp. 72-93. T.M. Scanlon, *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, MA: Belknap Press of Harvard University Press, 2008).

to eat your pizza. You prudently drop your pizza and back away, and the bear happily scampers off with it.

As in the previous scenarios, you will reasonably feel disappointed that you will not be able to eat your pizza. You may chastise yourself for not stowing it in a bear-proof container or in packaging that better masked its scent. But insofar as you do not make the mistake of personifying the bear, you will not resent it for taking your pizza, as you would if another person had intimidated you into giving it up. You realize that the bear is not an appropriate object for your norm-directed expectations, because it lacks the capacities to grant you the sorts of regard you demand from persons.<sup>105</sup>

You therefore adopt toward the bear what Strawson calls the *objective attitude*, suspending your normal range of reactive attitudes and thinking of it merely as a being to be understood, managed, or taken account of, not as a being of which certain attitudes can be appropriately demanded.<sup>106</sup> In a similar way, and for the same kinds of reasons, we sometimes adopt variations of the objective attitude toward other humans, such as young children and those with severe psychological deficiencies. As Gary Watson argues,

---

<sup>105</sup> Resenting the bear for taking your pizza would be like resenting the rain for ruining your picnic. The weather is not the sort of thing with which we can stand in a relationship of mutual accountability, and neither is the bear. A closely related point: in addition to the bear lacking the capacity to grant you moral standing, it lacks the capacity for *self-reactive attitudes*. These attitudes respond to one's norm-directed expectations for oneself, usually concerning the morality of one's behavior toward others. Within this category, Strawson mentions "such phenomena as feeling bound or obliged (the 'sense of obligation'); feeling compunction; feeling guilty or remorseful or at least responsible; and the more complicated phenomenon of shame." Strawson (2003), pp. 84-85.

<sup>106</sup> Korsgaard has pointed out in correspondence that there remains a philosophical puzzle here about how we view and interact with non-human animals. Our relationships with them can be affective and personal, even if they cannot be properly second-personal. We can grant an animal a kind of privileged standing—for instance, we take ourselves to be responsible for the wellbeing of beloved pets, and we would not normally treat them as potential sources of food. Moreover, we often want pets to reciprocate our affection in various ways. So our love or care for an animal can take us out of an objective stance and into a reciprocal relationship that is nevertheless not second-personal. Moreover, of course, we take ourselves to have moral obligations even to animals with which we do not have caring, reciprocal relations. One upshot of these points is that we are capable of taking up a wider variety of stances in our social interactions than the two stances Strawson canvasses.

To be intelligible, demanding presumes understanding on the part of the object of the demand. The reactive attitudes are incipiently forms of communication, which make sense only on the assumption that the other can comprehend the message.<sup>107</sup>

We do not hold young children and those with severe psychological deficiencies to the standards of normal interpersonal relationships because we recognize that they are not (yet) persons, and so are not capable of satisfying our norm-directed expectations in the ways that constitute competent participation in those relationships.<sup>108</sup>

Strawson argues that we can also sometimes adopt the objective attitude toward persons “as a refuge, say, from the strains of involvement; or as an aid to policy; or simply out of intellectual curiosity.”<sup>109</sup> But he thinks it would be “practically inconceivable” for us to adopt a globally objective attitude, since to do so would make it that case that “there were no longer any such things as interpersonal relationships as we normally understand them...[since these depend] precisely [on] being exposed to...reactive attitudes and feelings.”<sup>110</sup> Later, he argues that

[This] complicated web of attitudes and feelings...form[s] an essential part of the moral life as we know it...[and is] something we are given with the fact of human society...Our practices do not merely exploit our natures, they express them.”<sup>111</sup>

For Strawson, proneness to reactive attitudes is essential to personhood itself, since being prone to those attitudes is essential to having the capacity and disposition to participate in interpersonal relationships, and the latter is essential to being a person at all.

---

<sup>107</sup> Gary Watson, “Responsibility and the Limits of Evil: Variations on a Strawsonian Theme,” in Watson, *Agency and Answerability: Selected Essays* (New York: Oxford, 2004), p. 230.

<sup>108</sup> This was the point made in the third “exceptional” case in §4.2. Of course, we also grant young children and those with severe psychological deficiencies many of the same rights that we grant to persons, even though they cannot grant them back to us. As I noted in §2.2, a comprehensive MD theory of moral normativity (extending beyond the interpersonal dimension on which Intersubjectivism focuses) would need to account for this dimension of our moral lives.

<sup>109</sup> Strawson (2003), pp. 79-80.

<sup>110</sup> *Ibid.*, p. 81.

<sup>111</sup> *Ibid.*, pp. 91, 93.

In the next section, I consider three objections to my proposed account of personhood.

#### 4.4 Objections

First, one might object that I've merely stipulated my way to an interesting conclusion. I have characterized persons as beings whose drive for recognition motivates them to seek moral respect by participating in relationships of mutual accountability. But, of course, many theorists have thought that we can understand what persons are without making essential reference to morality or to our activities of holding one another accountable. For instance, Harry Frankfurt's influential account holds that one is a person in virtue of possessing a will with a distinctive structure: a person can form *second-order volitions*, second-order desires that certain of one's first-order desires be effective in moving one to act.<sup>112</sup> According to this view, a person is an agent who can reflect on his desires and who wants to act on some of them, but not others. The first objection, then, complains that I need to provide a good argument for why we should prefer my characterization of personhood to that of someone like Frankfurt.

The objection gains some of its apparent force from the thought that two theorists using the same term must be intending to refer to the same thing. This thought is true in one way and misleading in another, so my response is to partially concede and partially deflect the worry. First, the deflection: Frankfurt and I are focused on somewhat different things. I am using the word "person" to pick out those beings as we normally understand them, whom we see and interact with in everyday life and whose collective activities create our

---

<sup>112</sup> Harry Frankfurt, "Freedom of the Will and the Concept of Person," *Journal of Philosophy* 68 (January 1971), pp. 5-20. Reprinted in Frankfurt, *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), pp. 11-25.

shared social world. By contrast, Frankfurt is using the term to pick out the class of beings capable of a certain kind of reflective agency. I think that possessing the latter capacity is necessary but not sufficient for possessing the former, though I will not argue for that claim here. We could call the class of agents I am interested in by a different name—say, “social persons”—and the arguments I make would work just as well (or poorly). My claim is just that there is a certain kind of agent that we can only understand fully by appealing to its relationships of accountability with others of its kind. I call these agents “persons.”

Now, the concession: Frankfurt and I *are* focused on the same thing insofar as we are both aiming to elucidate the distinctive features of the type of agent we are. And, of course, whether one is classified as a person is very important, since judgments about personhood are normally prerequisites for judgments about obligation and accountability. So we need a good rationale for any proposed classification. I believe that my view that personhood is best understood in relational and moral terms gains support from this observation, since my view better comports with and makes sense of our ordinary practices than does Frankfurt’s.

A second objection holds that we could accept the conceptual connection between personhood and moral engagement but deny that we should understand either concept in terms of second-personal relations. After all, many have taken themselves to be bound by moral obligations without taking those obligations to be *owed to other persons* in virtue of the others’ standing. Such obligations have been thought to be owed to God, or to stem from independent, third-personal norms concerning the importance of wellbeing, or to be derived from considerations of self-interest. It would be implausible to hold that individuals who understand morality in one of these alternative ways are not persons, especially if they comply with more or less the same norms as does one who self-consciously takes up the second-personal standpoint discussed above.

I agree, but will mention two points of clarification in response. First, I think that these people are misinterpreting the nature of what they are up to, metaethically speaking. They feel the force of moral requirements but employ a mistaken conception of what their moral concepts are fundamentally about. If this claim counts as a bullet, then I am willing to bite it.

Second, and relatedly, the relationship we have with other persons is likely to be strained if we realize that they treat us as we think they ought to only because, say, they believe divine punishment would await them if they did not. We most naturally take such treatment to be owed *to us* in virtue of our standing to make claims on others. If another person fails to grant us the respect we think they owe us, we are disposed to respond with resentment and a demand for the appropriate recognition. If I am right about the essentially second-personal nature of moral interactions, then someone fails to fully exercise her capacity for personhood when she fails to grant others respect, e.g., by treating others as “counting” morally, but not as *owed* the relevant treatment.

The third objection picks up on this last point. If we understand personhood in terms of a *capacity* for certain kinds of second-personal relations, then it might seem an overstatement to claim that personhood is *essentially* second-personal. One could still count as a person if one possessed the capacity but never exercised it. Two types of case suggest themselves. The first involves someone who systematically fails to grant others respect, though (let’s assume) she has the capacity to do so. The second involves someone in a Robinson Crusoe-type scenario, who would (let’s assume) grant others respect if given the opportunity, but who never has the opportunity because she never encounters another person.

In the first case, I am willing to claim that someone who systematically refuses to participate in relationships of mutual accountability is failing to be a person. On my view, to classify some agent as a person is to make a judgment about how other persons can practically relate to her. Stated a bit differently, “person” is a metaphysical classification motivated by practical and epistemic concerns. We care about the capacity to participate in second-personal relationships because we care that the agent could be expected to (come to) exercise it in the appropriate circumstances. If this expectation is reasonable, then our attempts to hold the agent accountable are felicitous. If the expectation is unreasonable, then the agent is someone with whom we can only interact third-personally—say, by trying to avoid, control, or incentivize her in non-moral ways.

Of course, more would need to be said to fully specify the relevant counterfactuals. For instance, consider a deeply racist slave-owner in the antebellum South. How should his slaves conceive of their relation to him? Let’s assume that he engages in normal moral relations with whites, demonstrating that he clearly has the capacity to do the same with those whom he is enslaving. Let’s also assume that his racist judgments are so deeply entrenched that his slaves cannot reasonably expect that he will ever come to believe that he owes them certain kinds of moral consideration. There is then a sense in which he does and a sense in which he does not have the capacity to enter into relationships of mutual accountability with those whom he has enslaved. I suspect this realization would lead to a reasonable tension in the slaves’ attitudes about how they ought to try to relate to the slave-owner, i.e., whether to try to engage him as a person, or merely as a monster who has control over their lives.

With respect to less clear-cut cases, we would need to specify some minimum threshold (of likelihood of coming to exercise the capacity) and probably admit a fair degree

of vagueness when judging a range of borderline cases. This general challenge is shared by all views that appeal to a capacity as criterial for possessing some status, yet that also need to distinguish between the capacity and its regular exercise to explain what makes the thing a thing of the kind that it is.

In response to the Robinson Crusoe-type of case, I suggest that we distinguish between persons as we know them and persons as we might conceive of them. We might conceive of a person stranded alone on a desert island for decades who nevertheless retains the capacities to engage in second-personal relationships throughout the duration, such that she could step right back into normal relationships if she were then to encounter other persons. Unfortunately, persons as we know them—in the form of human beings—do not work like this. Solitary confinement, even in a tropical paradise, can be psychological torture for any social animal, humans included. If extended periods of social deprivation can degrade a human’s capacities for personhood, then we should conclude that personhood for human beings can be relational in a further sense: we need some degree of second-personal interaction in order to remain persons ourselves. Since the goal of Intersubjectivism is to understand how we can reconcile the views of *ourselves* from our practical and naturalistic perspectives, it is more pertinent to appeal to features of persons as we know them and not merely to features of persons as we might conceive of them.

### **5. Defending Intersubjectivism**

The moral psychological account constructed in §§3-4 enables us to respond to the initial worry that our intersubjective activities are too contingent to generate moral normativity. If Strawson and I are correct about the essential interconnection between being a person, being subject to the drive for respect, and participating in relationships of mutual

accountability, then it is not contingent that we engage one another morally. We could not fail to do so without abandoning our very natures. From one perspective, of course, to claim that moral normativity depends on our intersubjective activities is still to characterize it as contingent. Humans might have evolved very differently, or never at all, such that there might have never been such things as persons or relationships of mutual accountability. This kind of contingency is impossible to avoid for a MD view.

From another perspective—the perspective we take up when we think about whether our moral practices have a normative foundation—the fact that we *are* persons means that our intersubjective moral activities are practically necessary for us. Recognition of this practical necessity is a balm against practical worries about contingency.

Intersubjectivism holds that moral normativity does not require a non-contingent, mind-independent foundation. We can account for moral normativity by appealing to what we have contingently created, and what we continually maintain, through our thinking and acting together. This foundation is intersubjective, but the extent of its contingency is limited by the fact that our nature as persons disposes us to grant one another the practical authority that makes our moral interactions possible.

An objector might reply that there is an additional dimension to the worry about the contingency: our intersubjective activities are fallible. We have created many unjust social relationships over the course of human history. Since we owe these oppressive relationships to humans thinking and acting together, how could we expect the very same mechanisms to generate moral normativity? To put the point differently, we must leave conceptual space between moral normativity and whatever moral beliefs people happen to hold, even if they all happen to agree. Some group of people could, in principle, come to agree on nearly any set of moral beliefs. But the mere fact of their intersubjective agreement wouldn't make

those beliefs correct. Therefore, we can't treat moral normativity itself as a product of such agreement.

This is an important point, but we must take care not to overstate its relevance for the issue at hand. Consider three points. First, we don't come to metaethical inquiry to figure out which first-order moral beliefs we ought to hold. Accordingly, Intersubjectivism should not be interpreted as offering a reductive formula for making that first-order determination. It does not hold that whichever moral belief is intersubjectively supported is true. Second, and relatedly, we don't come to metaethical inquiry taking ourselves to be moral ignoramuses. We are confident in a great many of our first-order moral beliefs, and it is all but inconceivable that we would decide to change or abandon them upon further reflection (consider two classics: it is wrong murder innocent people, and it is good to alleviate avoidable suffering). What is the point of this inquiry, then? We experience the world as shot-through with moral normativity and we want a metaethical conception, if we can find one, that helps us make sense of that experience. We want to better understand how moral normativity could fit into our larger picture of the world, and we approach this question in part by reflecting on what we are doing when we make moral claims. Third, we needn't conceive of moral normativity as completely independent of our moral thought and activity in order to make room for the possibility of error. We are well aware of many ways in which our first-order moral reasoning can go astray. Even the saints among us will hold some mistaken moral beliefs due to factors like self-deception, subconscious selfishness, limited moral imagination, faulty inferences, or false empirical beliefs. In sum, the fact that our contingent intersubjective activities are fallible does not disqualify them from being creative with regard to moral normativity.

Recall that the third initial worry about Intersubjectivism holds that granting one another practical authority is insufficient to make it the case that we each *have* that authority. The natural thought here is that moral claims can be truly authoritative only if their normativity is objective. If there were no objective foundation for morality, then people could treat others however they wanted to and we'd have no way of rationally objecting. Any complaint in the face of purportedly immoral treatment would be a mere voicing of opinion or preference, or an expression of some non-cognitive attitude, or an attempt to control others through rhetoric or force. One might defend this thought by claiming that the objectivity of a norm is necessary for that norm's being *real*, in some important sense. This claim has been advocated not only by MI theorists, but also by defenders of Nihilism.

This approach to the question of moral authority is wrong-headed, because it ignores the essential practicality of moral claims. When you make a claim on me, what you presuppose is the existence of a certain kind of relationship between us, a certain standing you have in relation to me, and that I have in relation to you. It is our taking up this practical stance toward each other—our granting each other a distinctive power in our respective deliberative lives—that enables us create something that didn't exist before: a relationship of mutual accountability and reciprocal authority. On this view, moral authority is *co-authored*. It is a product of our relating to one another practically, rather than an already-given fact that provides us with reason to relate to one another in prescribed ways. Thus, according to Intersubjectivism, morality's authority is ultimately grounded in our relationships to *other people*, rather than (as many other views imply) ultimately grounded in our relationships to

*propositions* or to *states of affairs*.<sup>113</sup> I take this to be an appealing, if slightly revisionary, aspect of the view.

The case of friendship once again provides a useful parallel. There is no worry that friendship lacks a normative foundation just because it depends on our intersubjective thoughts and activity. It is our mutual, practical commitment to relating to others in certain ways—to abiding by the norms of friendship—that *makes* us friends to one another. Likewise, it is our mutual, practical commitment to relating to others in certain ways—to granting one another practical authority to make moral claims—that *makes* us persons to one another.

Here an objector could reply: unlike friendships, which are optional, the moral relationship obtains among all persons regardless of their choice to participate in it. That is to say, while friendships are contingent relationships between persons, morality establishes a necessary relation between us. Here we see an intersection of the worry about accounting for non-contingency and the worry about accounting for authority. We have already considered Intersubjectivism’s attempt to capture a version of this necessity claim, in the form of its argument that the very nature of persons disposes us to grant one another practical authority. The objector could agree that persons are *disposed* to relate to one another morally but put forward a separate objection. By attempting to ground moral authority in what

---

<sup>113</sup> Korsgaard’s view is similar in spirit, though often divergent in detail. She writes, “The normative demands of reason and meaning are not demands that are made on us by objects, but demands that we make on ourselves and each other.” Korsgaard (1996), p. 138. Elsewhere, she writes,

To say that you have a reason is to say something *relational*, something which implies the existence of another, at least another self. It announces that you have a claim on that other, or acknowledges her claim on you. For normative claims are not the claims of a metaphysical world of values upon us: they are claims we make on ourselves and each other. It is both the essence of consequentialism and the trouble with it that it treats The Good, rather than people, as the source of normative claims. The acknowledgment that another is a person is not exactly a reason to treat him in a certain way, but rather something that stands behind the very possibility of reasons.

Korsgaard, (1993), p. 51.

agents contingently decide to do—i.e., grant one another practical authority—

Intersubjectivism leaves open the possibility that relatively few persons will act on those dispositions, and therefore that relatively few will end up being morally bound to one another. If that possibility were realized, Intersubjectivism would account for moral authority only in a diminished sense, given that that authority would lack universal application. Furthermore, it is part of our common-sense moral understanding that we should be able to make moral claims even against those persons who fail to grant us practical authority. In fact, it is often when and because others fail to grant us authority that we need to make claims on them. But Intersubjectivism seems to rule out this possibility.

The first objection raises an empirical worry about the possible prevalence of systemic psychological misalignment. It might turn out that only a small number of us grant one another the practical authority that is constitutive of standing in the moral relationship together. If that were so, then there would be no common moral principles or shared understanding that we could appeal to in our dealings with one another, so the domain of morality's authority would be small. Here is an empirical refutation of this empirical objection: simply look around. We are surrounded by persons who relate to one other based on the shared understanding that each owes the other certain forms of respect.

Moreover, Intersubjectivism's appeal to the drive for respect can help explain why people who interact with one another tend, over time, to converge in their understandings of what morality most centrally requires—namely, because satisfying others' respect-seeking claims is one of the most effective ways of winning respect oneself. This convergence is crucial to Intersubjectivism's viability, since the theory relies on the idea that we are engaged in the same, shared practices of holding each other accountable. If these practices were not

shared, then our interaction would not amount to anything and morality's normativity could not have a genuinely intersubjective grounding.

The second objection argues that making moral normativity contingent and mind-dependent prevents us from being able to make moral claims on those who disrespect us. We can largely defuse this worry by taking a closer look at several scenarios in which another might fail to grant us respect. First, notice that most cases in which someone fails to grant another the (perceived) appropriate respect involve what we might call "local disrespect." Here, there is a wider context of granted respect, yet the disrespectful person withholds some more specific form of respect, often based on faulty reasoning or false empirical beliefs. For example, think of racist or sexist attitudes based on mistaken views about inherent intelligence. In these cases, one can in principle argue from true empirical claims and consistency with shared basic moral commitments to convince the other to grant the (perceived) appropriate respect. There may of course be many practical or irrational obstacles to the other's actually granting one the appropriate respect, but if they do fail in this way, they will be making a mistake by their own (shared) lights and will be appropriately blameworthy.

It is important that in defending Intersubjectivism we don't paper over the regularity with which people fail to live up to the standards of moral relationships. People treat each other badly very often. And the motivations that prevent us from treating one another even worse are often prudential, rather than moral. For instance, someone might refrain from stealing out of fear of being arrested, ostracized, or attacked, rather than out of respect for the person who owns the property. Intersubjectivism needn't deny these observations as it focuses on our intersubjective activities. It posits that our interpersonal relationships are mediated by acceptance of norms that ideally function as shared commitments. If someone

satisfies these commitments only half-heartedly, or intermittently, or for (what we think are) the wrong reasons, our recognition of these facts will affect what kind of relationship we can have with her, and will be reflected in the reactive attitudes and behaviors we form in response to her. The same holds for other kinds of interpersonal relationships, such as friendship. If we realized that someone was pretending to be our friend only to advance her self-interest, this would damage the relationship just as it would if we realized that someone was abiding by moral norms merely as a way of advancing or protecting his self-interest.

The second scenario in which another fails to grant one respect involves the other lacking the capacity to do so. In this scenario, the other is not a person in the sense we are concerned with. One will appropriately interact with the agent from Strawson's objective stance, given that it is a condition of the felicity of moral address that one takes the other to have the capacity to respond appropriately to one's address. Finally, in the third scenario, the other is a young child who does not yet possess the fully functioning capacity for moral interaction, but who could develop that capacity with the right guidance. In this case, we might make moral claims on the child as a practical way of teaching the child how to become a person. The upshot of this discussion is that in none of these scenarios do others' failures to express respect undermine Intersubjectivism.

The third scenario just canvassed provides a useful segue to discuss the multiple functions of making a moral claim. The primary function of the activity is to hold another accountable to the standards that govern one's relationship with the other. This function is essentially practical—we exhort another to (continue to) stand in a moral relationship with us and to act in the ways called for by the norms of that relationship and by our respective standings within it. Therefore, in making a moral claim, we both call upon and express the nature of our relationship with another. In response to the third scenario above, we can

think of moral claim-making as an attempt to establish such a relationship. One can attempt to establish a moral relationship by affecting which aspects of the world the other is sensitive to and thus which claims the other is disposed to grant consideration. Neither of these practical functions of making moral claims is incompatible with the additional function of describing what morality requires, either by stating shared moral commitments or by arguing for what follows from those commitments. There is a sort of “correspondence” here, though it is with a domain of co-created reality, i.e., the norms governing our relationships of mutual respect. The great majority of moral claims play both functions at once.

## **6. Conclusion**

My aim in this chapter has been to introduce Intersubjectivism and to argue that it is a promising candidate for solving the reconciliation problem. Much of this effort has involved teasing out interconnections between concepts that structure how we conceive of ourselves and of our relations to others. One of my goals has been to show that the objects of these conceptions—our selves and our relations to others—are more intimately intertwined than we normally acknowledge. I have argued that persons are essentially second-persons, and that this realization makes it possible to conceive of moral normativity as essentially intersubjective. While this central claim of Intersubjectivism will have struck many as (at least) initially implausible, my goal throughout has been to construct a picture of our moral lives in which we can recognize ourselves: how we think, what we care about, and how we interact with one another. And I have tried to capture what seems to me so special and surprising about morality itself. The fact that we can relate to one another in these ways is a hard-won achievement, which we constantly renew together.

## -BIBLIOGRAPHY-

- Alston, William. "Epistemic Circularity." *Philosophy and Phenomenological Research* 47, 1986.
- . *The Reliability of Sense Perception*. Cornell University Press, 1993.
- Aristotle, *Nicomachean Ethics*. Translated by Christopher Rowe, with commentary by Sarah Broadie. Oxford University Press, 2002.
- Barnett, David J. "What's the Matter with Epistemic Circularity?" *Philosophical Studies*, Vol. 171, No. 2, 2014.
- Bengson, John, Terence Cuneo, and Russ Schafer-Landau. *Understanding Morality: Method, Content, and Essence*. Oxford University Press, forthcoming.
- Bergmann, Michael. "Externalism and Skepticism." *Philosophical Review* 109, 2000.
- . "Epistemic Circularity: Malignant and Benign." *Philosophy and Phenomenological Research* 69, 2004.
- . "Externalist Responses to Skepticism." *Oxford Handbook of Skepticism*. Edited by John Greco. Oxford University Press, 2008.
- Berker, Selim. "Does Evolutionary Psychology Show that Normativity is Mind-Dependent?" *Moral Psychology and Human Agency: Essays on the New Science of Ethics*. Edited by Justin D'Arms and Daniel Jacobson. Oxford University Press, 2015.
- Blackburn, Simon. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford University Press, 1984.
- . *Essays in Quasi-Realism*. Oxford University Press, 1993.
- . *Ruling Passions: A Theory of Practical Reasoning*. Clarendon Press, 2001.
- Chisholm, Roderick M. *The Foundations of Knowing*. The University of Minnesota Press, 1982.
- C.H. Cooley, *Human Nature and the Social Order*. Scribner's, 1902.
- Cuneo, Terence. *The Normative Web: An Argument for Normative Realism*. Oxford, 2007.
- Cuneo, Terence, and Russ Shafer-Landau. "The Moral Fixed Points: New Directions for Moral Nonnaturalism." *Philosophical Studies* 171, 2014.
- Darwall, Stephen. "Two Kinds of Respect." *Ethics* 88, 1977.

- . *The Second-Person Standpoint: Morality, Respect, and Accountability*. Harvard University Press, 2006.
- . *Morality, Authority, and Law: Essays in Second-Personal Ethics I*. Oxford University Press, 2013a.
- . *Honor, History, and Relationship: Essays in Second-Personal Ethics II*. Oxford University Press, 2013b.
- Dawkins, Richard. *The Selfish Gene*. Oxford University Press, 1976.
- . *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution*. Houghton Mifflin, 2004.
- Dent, N. J. H. *Rousseau: An Introduction to his Psychological, Social, and Political Theory*. Blackwell, 1989.
- Dworkin, Ronald. "Truth and Objectivity: You'd Better Believe it." *Philosophy and Public Affairs*, Vol. 25, No. 2, Spring 1996, pp. 87-139.
- . *Justice for Hedgehogs*. The Belknap Press of Harvard University Press, 2011.
- Eilan, Naomi. "The You Turn." *Philosophical Explorations*, 2014.
- Enoch, David. "How Objectivity Matters." *Oxford Studies in Metaethics*, Vol. 5, 2010.
- Enoch, David, and Joshua Schechter. "How are Basic Belief-Forming Methods Justified?" *Philosophy and Phenomenological Research*, Vol. 76, No. 3, 2008.
- Frank, Robert H. *Passions within Reason: The Strategic Role of the Emotions*. Norton, 1988.
- Frankfurt, Harry. "Freedom of the Will and the Concept of Person." *Journal of Philosophy* 68 (January 1971). Reprinted in Frankfurt, *The Importance of What We Care About*. Cambridge University Press, 1988).
- Gibbard, Allan. *Wise Choices, Apt Feelings: A Theory of Moral Judgment*. Harvard University Press, 1992.
- . *Thinking How to Live*. Harvard University Press, 2008.
- . *Reconciling Our Aims: In Search of Bases of Ethics*. Oxford University Press, 2011.
- Goldman, Alvin I. "What is Justified Belief?" in G.S. Pappas (ed.) *Justification and Knowledge* (Dordrecht: Reidel, 1979), pp. 1-25.
- Harman, Gilbert. *The Nature of Morality: An Introduction*. Oxford University Press, 1977.
- Harman, Gilbert, and Judith Jarvis Thomson. *Moral Relativism and Moral Objectivity*.

- Wiley Blackwell, 1996.
- Honneth, Axel. *The Struggle for Recognition: The Moral Grammar of Social Conflicts*. MIT, 1996.
- Huemer, Michael. *Ethical Intuitionism*. Palgrave Macmillan, 2005.
- . "Compassionate Phenomenal Conservatism." *Philosophy and Phenomenological Research* 74, 2007.
- . "Phenomenal Conservatism." *Internet Encyclopedia of Philosophy*.  
<http://www.iep.utm.edu/phen-con/print#SH3d>.
- Horgan, Terry, and Mark Timmons. "What Does Moral Phenomenology Tell us about Moral Objectivity?" *Social Philosophy and Policy* Vol. 25, No. 1, 2008.
- Joyce, Richard. *The Evolution of Morality*. MIT Press, 2006.
- . *The Myth of Morality*. Cambridge University Press, 2007.
- . "The Evolutionary Debunking of Morality." *Reason and Responsibility*. Edited by J. Feinberg and R. Shafer-Landau. Cengage, 2013.
- . "Evolution, Truth-tracking, and Moral Skepticism." *Essays in Moral Skepticism*. Oxford University Press, 2016.
- . "Human Morality: From an Empirical Puzzle to a Metaethical Puzzle." *Cambridge Handbook of Evolutionary Ethics*. Edited by M. Ruse and R. Richards. Cambridge University Press, forthcoming.
- Kant, Immanuel. *The Metaphysics of Morals*. Translated by Mary Gregor. Cambridge University Press, 1996.
- . *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor. Cambridge University Press, 1998.
- Kolodny, Niko. "Objectivity in Ethics." *Encyclopedia of Philosophy, Second Edition*. Edited by Donald Borchert. Macmillan, 2006.
- . "The Explanation of *Amour-Propre*." *Philosophical Review*, 119:2, 2010.
- Korsgaard, Christine M. *The Sources of Normativity*. Cambridge University Press, 1996.
- . *Creating the Kingdom of Ends*. Cambridge University Press, 1996.
- . *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford University Press, 2008.

- . *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press, 2009.
- Lavin, Douglas. "Other Wills: The Second-Person in Ethics." *Philosophical Explorations*, Vol. 17, No. 3, 2014.
- Lammenranta, Markus. "Epistemic Circularity." *Internet Encyclopedia of Philosophy*.  
<http://www.iep.utm.edu/ep-circ/>
- Lenman, James, and Yonatan Shemmer, editors. *Constructivism in Practical Philosophy*. Oxford University Press, 2012.
- Lovejoy, A.O. *Reflections on Human Nature*. The Johns Hopkins Press, 1961.
- Lynch, Michael P. and Paul Silva, Jr. "Why Worry About Epistemic Circularity?" *Journal of Philosophical Research: Supplement—Papers in Honor of William P. Alston*, 2016.
- Mackie, J.L. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.
- McPherson, Tristram. "Metaethics and the Autonomy of Morality." *Philosophers' Imprint*, Vol. 8, No. 6, 2009.
- . "Against Quietist Normative Realism." *Philosophical Studies*, Vol. 154, No. 2, 2011.
- Mead, G.H. *Mind, Self, & Society: From the Standpoint of a Social Behaviorist*. Edited by Charles W. Morris. The University of Chicago Press, 1934.
- Nagel, Thomas. *The View from Nowhere*. Oxford University Press, 1989.
- . *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. Oxford University Press, 2012.
- Neuhouser, Fredrich. *Rousseau's Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition*. Oxford University Press, 2008.
- Nietzsche, Friedrich. *On the Genealogy of Morality*. Translated by Maudemarie Clark and Alan J. Swensen. Hackett Publishing, 1998.
- Nozick, Robert. *Anarchy, State, and Utopia*. Basic Books, 1974.
- . *Philosophical Explanations*. Harvard University Press, 1981.
- Parfit, *Reasons and Persons*. Oxford University Press, 1984.
- . *On What Matters: Volume 1*. Oxford University Press, 2011a.
- . *On What Matters: Volume 2*. Oxford University Press, 2011b.
- Pascal, Blaise. *Pensées*. Translated by A.J. Krailsheimer. Penguin Classics, 1995.

- Piaget, Jean. *The Origins of Intelligence in Children*. Basic Books, 1952.
- . *The Construction of Reality in the Child*. Basic Books, 1954.
- Rawls, John. "Kantian Constructivism in Moral Theory." *Journal of Philosophy* 77, 1980.
- . *A Theory of Justice, Revised Edition*. The Belknap Press of Harvard University Press, 1999.
- Rochat, Philippe. *Others in Mind: Social Origins of Self-consciousness*. Cambridge University Press, 2009.
- Rousseau, Jean-Jacques. *Discourse on the Origin and Foundations of Inequality among Men, or Second Discourse*. In *Rousseau: The Discourses and Other Early Political Writings*. Translated and edited by Victor Gourevitch. Cambridge University Press, 1997.
- . *Emile, or on Education*. Translated by Allan Bloom. Basic Books, 1979.
- Scanlon, T.M. *What We Owe to Each Other*. The Belknap Press of Harvard University Press, 1998.
- . *Moral Dimensions: Permissibility, Meaning, Blame*. The Belknap Press of Harvard University Press, 2008.
- . *Being Realistic about Reasons*. Oxford University Press, 2014.
- Schmidt, Frederick F. "What is Wrong with Epistemic Circularity?" *Philosophical Issues* 14, 2004.
- Sinnott-Armstrong, Walter. *Moral Skepticisms*. Oxford University Press, 2007.
- Skyrms, Brian. *Evolution of the Social Contract*. Cambridge University Press, 1996.
- Smith, Adam. *Lectures on Jurisprudence*. Edited by R.L. Meek, et al. Liberty Classics, 1982.
- . *The Theory of Moral Sentiments*. Edited by D.D. Raphael and A.L. Macfie. Liberty, 1985.
- Steup, Matthias. "Is Epistemic Circularity Bad?" *Res Philosophica* 90, 2013.
- Strawson, Peter. "Freedom and Resentment." *Proceedings of the British Academy* 48, 1962.
- Street, Sharon. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127, 2006.
- . "Constructivism about Reasons." *Oxford Studies in Metaethics*, Vol. 3, 2008.

- . "In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters." *Philosophical Issues*, Vol. 19, 2009a.
- . "Evolution and the Normativity of Epistemic Reasons." *Canadian Journal of Philosophy*, Vol. 39, 2009b.
- . "What is Constructivism in Ethics and Metaethics." *Philosophy Compass*, Vol. 5, No. 5, 2010.
- . "Mind-Independence without the Mystery: Why Quasi-Realists Can't Have it Both Ways." *Oxford Studies in Metaethics*, Vol. 6, 2011.
- . "Coming to Terms with Contingency: Humean Constructivism about Practical Reason." *Constructivism in Practical Philosophy*. Edited by James Lenman and Yonatan Shemmer. Oxford University Press, 2012.
- . "Does Anything Really Matter or Did We Just Evolve to Think So?" *The Norton Introduction to Philosophy*. Edited by Alex Byrne, et al. Norton, 2015.
- Thompson, Michael. "What is it to Wrong Someone? A Puzzle about Justice." *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Edited by R. Jay Wallace, et al. Oxford University Press, 2004.
- Tomasello, Michael. "The Interpersonal Origins of Self-Concept." *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*. Edited by U. Neisser. Cambridge University Press, 1993.
- . *The Cultural Origins of Human Cognition*. Harvard University Press, 1999.
- . *Why We Cooperate*. The MIT Press, 2009.
- Vavova, Katia. "Debunking Evolutionary Debunking." *Oxford Studies in Metaethics*. Edited by Russ Shafer-Landau, Vol. 9, 2014.
- . "Irrelevant Influences." *Philosophy and Phenomenological Research*, 2016.
- Watson, Gary. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." *Agency and Answerability: Selected Essays*. Oxford, 2004.
- Weisberg, Jonathan. "Bootstrapping in General." *Philosophy and Phenomenological Research*, Vol. 81, No. 3, 2010.
- . "The Bootstrapping Problem." *Philosophy Compass*, Vol. 7, No. 9, 2012.
- Wiggins, David. "Objectivity in Ethics: Two Difficulties, Two Responses." *Ratio*, Vol. 18, 2005.

Wolf, Susan. "Sanity and the Metaphysics of Responsibility." *Responsibility, Character, and the Emotions*. Cambridge University Press, 1987.

Wright, Crispin. *Truth and Objectivity*. Harvard University Press, 1994.