9/03

Note to readers of *The Terrible, Horrible, No Good, Very Bad Truth about Morality and What to Do About it*, Doctoral Dissertation of Joshua D. Greene in the Department of Philosophy, Princeton University, June 2002.

You are welcome to read this work, pass it on to others, and cite it. I only ask that if you pass on this work to someone else that it be passed on (1) in its entirety, (2) without modification, and (3) along with this note.

I consider this a work in progress. It is currently under review in its present form at an academic press. I intend to revise and expand it substantially before publishing it as a book, so much so that the book and the dissertation will probably best be considered separate works.

Comments are welcome. You can contact me by email (jdgreene@princeton.edu) or by regular mail:

Joshua Greene
Department of Psychology
Princeton University
Princeton, NJ 08544

jdg

THE TERRIBLE, HORRIBLE, NO GOOD, VERY BAD TRUTH ABOUT

MORALITY AND WHAT TO DO ABOUT IT

Joshua David Greene

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT

OF PHILOSOPHY

NOVEMBER 2002

Abstract

In this essay I argue that ordinary moral thought and language is, while very natural, highly counterproductive and that as a result we would be wise to change the way we think and talk about moral matters. First, I argue on metaphysical grounds against moral realism, the view according to which there are first order moral truths. Second, I draw on principles of moral psychology, cognitive science, and evolutionary theory to explain why moral realism appears to be true even though it is not. I then argue, based on the picture of moral psychology developed herein, that realist moral language and thought promotes misunderstanding and exacerbates conflict. I consider a number of standard views concerning the practical implications of moral anti-realism and reject them. I then sketch and defend a set of alternative revisionist proposals for improving moral discourse, chief among them the elimination of realist moral language, especially deontological language, and the promotion of an anti-realist utilitarian framework for discussing moral issues of public concern. I emphasize the importance of revising our moral practices, suggesting that our entrenched modes of moral thought may be responsible for our failure to solve a number of global social problems.

in fond memory of

David Lewis

1941-2001

a philosopher without counterpart

in this world or any other

Acknowledgements

My philosophical interests are rather pedestrian as philosophical interests go. If your man or woman on the street takes no interest in a certain issue, chances are I don't either. One consequence of my distaste for esoterica is that my friends and family outside of philosophy are, to me, as valuable a professional resource as my colleagues in the field. As my ideas have developed I've tried to make them accessible and compelling to intelligent lay people. What this means in practice is that my non-philosopher friends and family have endured years of my ranting, and I thank them for their trouble—for listening, for keeping me grounded with their feedback (sometimes in the form of glazed expressions), and for providing me with what is loosely called "moral support." I thank my housemates Giulio Boccaletti and Antoinette Handley; my old friends Koichi Kurisu, Jon Bresman, and Darin McKeever; and my *old* old friends Brett Halsey and Emily Sobel. Thanks to Jessica Freireich for her open mind and heart. Thanks to Andrea Heberlein for a lifetime of affection in less than a year, with the best yet to come. Thanks to my inspiring siblings, Danny and Elizabeth, and to my parents, Laurie and Jonathan Greene, who taught me from the start to think for myself.

As an undergraduate, I was fortunate to have had three talented and generous mentors who nurtured me by taking my fledgling ideas seriously when others would have laughed at them. I am forever grateful to Jonathan Baron, Amartya Sen, and Derek Parfit for sending me on my way.

Much of my work in recent years has been outside of philosophy in the emerging field of social cognitive neuroscience, and I am very grateful for my second home in Princeton's psychology department. The influence of my work in psychology on my philosophical work is obvious, and my sincerest thanks go to Jonathan Cohen, John Darley, and Leigh Nystrom for taking me in and for providing me with unprecedented opportunities.

I would also like to acknowledge Judith Viorst, whose charming tale of pre-adolescent angst inspired the title of this work.[1]

Over the last five years I've had the privilege of defending my ideas against some of the best minds in the philosophy business, both young and old. My debts to the members of Princeton's philosophy department are many, and I will not attempt to enumerate my creditors for fear of omissions. A few, however, deserve special mention. Thanks to Caspar Hare and Simon Keller for their constant companionship and insight. Thanks also to Anja Jauernig, Angelique Knapp, Josh Knobe, Brian Lee, Jonathan and Phillippa McKeown-Green, Jessica Moss, Casey O'Callaghan, Kieran Setiya, and Jeff Speaks. Many thanks to Mark Johnston, Paul Benacerraf, and my peers who participated in our departmental dissertation seminars during my post-generals years. I am also grateful to Gideon Rosen and the students in his meta-ethics seminar (Spring 2001) for allowing me to present to them an earlier version of my second chapter and for providing me with valuable feedback. Thanks to Ann Getson for being so much more than her title, "Graduate Assistant," could ever convey.

---

[1] *Alexander and the Terrible, Horrible, No Good, Very Bad Day* (1987).

Table of Contents

Part I: The Terrible, Horrible Truth About Morality

Part II: What to Do About It

Part I:

The Terrible, Horrible Truth About Morality

Chapter 1

Introduction

"The unleashing of power of the atom bomb has changed everything except our mode of thinking, and thus we head toward unparalleled catastrophes."[1]

—Albert Einstein

This essay is an attack on common sense—moral common sense, in particular. Mounting evidence suggests that our sense[2] of right and wrong is a finely honed product of natural selection (Wright, 1994). We think about moral matters as we do in large part because our kind of moral thinking, in the heads of our prehistoric ancestors, enabled them to reproduce more effectively than their competitors, leaving us, their descendants, to inherit their world. But the world they left us is radically different from the world we now inhabit, and, as a result, what was biologically advantageous for them may prove disastrous for us. At the risk of

---

[1]Quoted in the *New York Times Magazine*, August 2, 1964 and in Calaprice (2000).

[2] As I will explain later (Chapter 3), the human moral sense isn't so much a single sense as it is a single capacity for acquiring many different moral senses, and these senses aren't so much senses as they are capacities for affective response.

being overly dramatic, I propose that the fate of humankind will turn on our ability, or inability, to transcend the common sense morality we inherited from our ancestors.   The great global problems of our time—the proliferation of weapons of mass destruction, the disruption of the environment, etc.—can only be solved through cooperation and compromise among people with radically different moral outlooks. And this, I believe, is unlikely to happen so long as the people of the world hold fast to their respective versions of moral common sense.

What is so terrible about moral common sense? In a phrase, *moral realism*. Moral realism is, roughly, the idea that there is a fact of the matter about what's right or what's wrong.[3] Suppose, for example, that you turn the corner to find a group of youngsters lighting a cat on fire for fun.[4] You are shocked and appalled and have a thought that may be expressed by the words: "That's wrong!" What does it mean for this action to be *wrong*? Is it wrong because you

---

[3] Some philosophers, chief among them Simon Blackburn (1984, 1993a, 1993b), have argued that the difference between realism and anti-realism in ethics may turn out to be obscure and of little consequence. I disagree. I believe that the question "Do moral questions ever have correct answers?" is rather clear and comprehensible as philosophical questions go, even to non-philosophers, and that the answers one may give to this question, a realist "yes" or an anti-realist "no," are distinct and meaningful. At this point, *pace* Blackburn, I will trust that the reader will glean my intended meaning of "moral realism." We will return to this issue shortly in Section 1.1.

[4] This example is borrowed from Gilbert Harman (1977).

dislike it or disapprove of it? You may have such negative attitudes toward this

behavior, but if you're an ordinary moralist you don't think that your having

these attitudes is what *makes* this behavior wrong. Is this behavior wrong

because the community to which you belong dislikes, disapproves of, or prohibits

it? Presumably not. If, somehow, the whole community were to decide that

torturing cats is okay, this activity would, by the lights of most people, still be

wrong. In answering these questions in the negative, one expresses an implicit

theory concerning the metaphysics of morals. Moral claims, some of them

anyway, are not merely *subjective*, dependent for their validity on the judgments

of individuals or groups of individuals, but rather *objective*. The fact that torturing

animals is wrong doesn't depend on what people think about it, although we may

hope that what people think about it has something to do with the fact that it's

wrong.

  And what is so terrible about thinking that some actions are objectively

wrong? Isn't it *good* that we take torturing animals to be *objectively wrong*? I

agree with common sense moralists that in this case, and others like it, the

attribution of objective wrongness is harmless. However, a general tendency to

think in terms of objective right and wrong is, I believe, quite problematic for two

reasons. First, and less important, it's simply mistaken. Contrary to common

sense, moral realism is false. In Chapter 2 I will attempt to explain why it's false,

and in Chapter 3 I will draw on recent work in moral psychology to explain why it

nevertheless strikes most people as true. Second, and more important, I

propose that moral realism makes it difficult for people with different values to get

along with one another because people who have practical disagreements are less willing to compromise if each of them thinks she has the Moral Truth on her side. In Chapter 4 I will explain in more detail how moral realism, given the nature of moral psychology and the present state of the world, gets us into trouble. I will then advance my proposal for avoiding this trouble, a view I call *revisionism*, by which I advocate the elimination of moral realism from our moral thought and language. Finally, In Chapter 5 I will explore some of the practical implications and intricacies of revisionism, including its implications for moral philosophy, moral discourse, and the resolution of moral conflicts great and small.

In this introductory chapter I will attempt to clarify the meta-ethical issue surrounding moral realism and provide some recent historical context.

1.1 Projectivist Anti-Realism

John Mackie (1977) argued that there are no objective values. He also claimed that ordinary moral language and thought presuppose the existence of objective values, and for this reason he characterized his meta-ethical view as an *error theory*. According to Mackie, claims such as "Lying is wrong" are, contrary to popular presupposition, never true.

What exactly Mackie meant in denying the existence of objective values is open to interpretation, but there is a broad consensus among Mackie's

interpreters that his denial of objective values makes him, among other things, a moral *anti-realist*. As a first pass, we can say that a moral anti-realist is one who believes that there are, to use Mackie's term, no *first order* moral truths or moral facts. We'll say that a first order moral claim is one that predicates (or denies the predication of) some moral property (e.g. wrongness) of some set of particular actions, persons, institutions, etc., where these particulars may be actual or merely possible. You might say that first order moral claims are ethical, as opposed to meta-ethical. Typical examples include claims such as "Lying is wrong," and, "The affluent ought to give to charity," and do not include claims such as "There are no moral facts."[5] (Throughout this essay, talk of "moral

_____

[5] While the distinction between first order and second order moral claims is no more exotic than the familiar distinction between ethics and meta-ethics, it is a bit fuzzy. For example, the claim that "No acts whatsoever are wrong" can be read as an exceptionally broad first order/ethical claim or as a standard anti-realist meta-ethical claim. Perhaps the best way to distinguish between first and second order claims is to say that only first order claims are action-guiding in any kind of straightforward way. To say that "Lying is wrong," is to speak against lying. Likewise, to make the first order claim that "No acts whatsoever are wrong," is to speak in favor of a kind of broad permissiveness. In contrast, to use these words to express a meta-ethical point is to make a broad claim about the nature of morality that may or may not have significant practical implications.

At present, my task is to distinguish anti-realism from realism. To this end, one might say that anti-realism is the view according to which all the moral truths

judgments" or "moral claims" should be understood as talk of first order moral

judgments/claims unless otherwise noted.) Mackie described the error embodied

in ordinary moral thought as one of presupposing that there are objective values

when in fact there are none. For our purposes, we can characterize this putative

error as one of presupposing that moral realism is true when in fact it is false.

Anti-realists such as Mackie do not deny that there are moral truths in the

broader sense of "truths concerning morality." Indeed, the claim that there are no

first order moral truths is one such truth. Nor do anti-realists deny that there are

truths concerning morality *qua* psychological and sociological phenomena. "Pro-

lifers" believe that abortion is wrong. The truth that they have this belief is a truth

in good standing and, in some sense, a moral truth, but this, too, is not the sort of

putative truth the anti-realist denies, as it is a truth concerning what people

believe about right and wrong and not a truth concerning which things are right

and wrong *per se*.

The above account of anti-realism requires some qualification, due to the

emergence of Simon Blackburn's (1984, 1993a, 1993b) *quasi-realism* and

"minimalist" approaches to moral truth.[6]   I will consider Blackburn's views in

greater detail in Section 1.4.3 and in Chapter 4, but a few words here will ward

off confusion in the meantime. Blackburn calls himself a moral anti-realist, but he

thinks that even a moral anti-realist can "earn the right" to talk like a moral realist,

---

are those that are true entirely in virtue of the non-existence of moral properties,

properties such as wrongness etc.

[6] For a recent example see Timmons (1999).

that is to use the propositional language of ordinary, first order moral discourse.

At the end of the day, says Blackburn, a savvy anti-realist is entitled to believe

that claims such as "lying is wrong" can be true, state facts, deserve to be called

"objective," etc. Given Blackburn's status as an anti-realist, there is a prima facie

problem with identifying anti-realists as those who deny that there are first order

moral truths/facts.

But if an anti-realist need not deny that there are moral truths/facts, what

does it take to be an anti-realist? To understand Blackburn's answer to this

question, we must consider the broadly Humean (Hume, 1888) notion of

projectivism. According to the moral projectivist, "We have sentiments or other

reactions caused by natural features of things and we 'gild or stain' the world by

describing it as if it contained features answering to these sentiments, in the way

that the niceness of an ice cream answers to the pleasure it gives us,"

(Blackburn, 1993a, pg. 152). In *Spreading the Word* (1984, pg. 171) Blackburn

writes:

> Suppose that we say we project an attitude or habit or other commitment
> which is not descriptive onto the world, when we speak and think as
> though there were a property of things which our sayings describe, which
> we can reason about, know about, be wrong about, and so on.

Blackburn is a moral projectivist, and he explains his anti-realism in terms of his

projectivism. (While projectivism and anti-realism are closely allied, it is not a

forgone conclusion that projectivism implies anti-realism. See Chapter 3.)

According to him, what makes him an anti-realist is not his denial of moral

facts/truths—Blackburn will ultimately affirm these things—but rather his belief

that moral judgments "start theoretical life" as "a stance, or conative state or

pressure on choice or action," (1993a, pg. 168).

I believe that Mackie and Blackburn have roughly the same meta-ethical

picture, though their respective choices of word make this obscure: Blackburn

says that there are a great many first order moral truths, while Mackie says that

no such claims are true. Mackie says that ordinary moral thought and language

are infected with error while Blackburn denies that this is so. Nevertheless, I

claim—and many philosophers would no doubt agree—that Blackburn and

Mackie are not really all that far apart, and this is reflected in the fact that they

both endorse *projectivist anti-realism*.

I find Blackburn's vocabulary confusing and misleading. Blackburn is an

anti-realist, which means that he doesn't think that moral properties are *real*, but

he is nevertheless willing to say that there are *truths* and *facts* concerning which

things have these unreal properties. Blackburn (1984, pg. 171) says that moral

judgments are not descriptive, but at the same time he wants to speak like a

realist, to say, for example, that "wrong" is an apt description of Nixon's behavior.

To keep things straight, I will speak in a non-Blackburnian fashion. Moral truths

and moral facts, as I employ these concepts, must not in general[7] depend[8] on

---

[7] Some *bona fide* moral facts (i.e. not quasi-facts) might be mind-dependent in

the following way. Suppose that it is wrong to do something that one believes to

9

be wrong. Were that so, the wrongness of a given act would depend on the agent's moral sensibilities and, more specifically, on his beliefs about the wrongness of the act. Hence the "in general" qualification above.

[8] One might wonder what one means by "depend" in this case, and a Blackburnian interested in trying to make anti-realism look as much like common sense as possible might deny that, according to his view, moral truths "depend" on our patterns of projection or moral sensibilities. I maintain, however, that in the end projectivism commits one to such a dependence.

At the very least the projectivist is committed to a causal story about our moral judgments. Our judgments concerning what things are right and wrong exhibit a causal dependence on our patterns of projection. According to Blackburn, the fact that we so often get things "correct" is not some coincidence between our patterns of projection and some set of further, moral facts. Rather, we know what's right and what's wrong by "perceiving" the properties we've projected onto the world. One might—and I think one should—then conclude that according to this view what's right and what's wrong *depends* on what our moral sensibilities happen to be. If our patterns of projection cause our moral judgments to be what they are (and we're not just getting lucky when we get things right), then it would seem that the truth of these judgments depends on our patterns of projection.   But Blackburn is not happy with this because it makes his view look like something ordinary people would not believe because most people don't think that the facts about what's right or wrong depend, in general, on what people think is right or wrong. What's a quasi-realist to do?

our patterns of projection, on our moral sensibilities.[9] If, for example, it's a *fact* that lying is wrong, if it's *true* that lying is wrong, this fact/truth cannot depend on

---

Out of the analytic philosopher's tool box comes the rigid designator. We can say that moral terms implicitly refer to the sensibilities that we happen to have in the actual world. Thus, when we consider a possible world in which people have different moral sensibilities, ones that we find abhorrent, we can still say that the things we think of as wrong are wrong in their world, too, because the content of "wrong" is determined by our sensibilities and not by theirs. See Davies and Humberstone (1980).

We will discuss this strategy in greater detail in Section 2.4.3, but for now I'll simply note that this strategy only pushes the bump in the carpet. Taking the irksome dependence of moral truth on our actual moral sensibilities and packing it into the meanings of moral terms gets rid of the straightforward counterfactual dependence discussed above, but only by introducing a new, equally irksome kind of dependence. Making it true *by definition* that first order moral truths depend on our actual moral sensibilities is no improvement. The problem with both kinds of dependence may be explained in terms of common sense morality's commitment to what I call the *ultimate asymmetry* between the things we call "right" and the things we call "wrong." (Again, See Section 2.4.3.)

[9] Others besides Blackburn may object to this characterization of realism. John McDowell (1985), for example, considers himself a realist through and through, but nevertheless claims that moral facts depend on our moral sensibilities in more or less the way that I've claimed that realism forbids. For the purposes of

our tendency to project wrongness onto acts of lying. When considering such

mind-dependent "truths, " "facts," "properties," and "realities" a la Blackburn I will

this essay, however, it makes more sense to classify self-described realists such

as McDowell as anti-realists of a Blackburnian stripe. In the next chapter I will

argue against moral realism as defined above. Some of the arguments and

comments from that chapter will apply to views such as McDowell's, but those

views are not the primary targets in that discussion. In Section 4.6.1 I will argue

against Blackburn's position on pragmatic grounds, and those arguments will

apply to views such as McDowell's as well.

It's worth noting that many of the views I classify as realist and discuss as

such in the next chapter could in principle be excluded from the realist camp, as

these are often views according to which the moral facts depend in some way on

our subjective responses. They are, however, typically views according to which

the moral facts are mind-dependent in a less straightforward way than they are

for McDowell and Blackburn. McDowell and Blackburn essentially endorse our

moral sensibilities as they are (though they make provisions for the possibility of

improving or correcting them), and insofar as they are realists (quasi- or

otherwise), they see moral inquiry as discontinuous with scientific inquiry

(Darwall, Gibbard, and Railton,1992). In contrast, the realists discussed in the

next chapter tend to view moral inquiry as continuous with scientific inquiry, as

exploring a domain of facts *a la* scientific facts. They tend to offer meta-ethical

theories that they hope will pave the way for a better, and potentially surprising,

understanding of what the moral facts are.

speak instead of "quasi-truths," "quasi-facts," "quasi-realities," "quasi-properties," etc. When I speak of anti-realism's being true or realism's being false, I am talking about the possibility that there are no moral truths/facts while leaving it open as to whether or not there are any moral quasi-truths or quasi-facts. When I speak of something's being right (or wrong), I speak of its being truly and factually right (or wrong) and not quasi-right (or quasi-wrong).

Mark Timmons' (1999) meta-ethical views are similar to Blackburn's, and he takes a similarly liberal approach to issues concerning what sorts of claims may legitimately be called "true." For truth-minimalists such as Timmons, to assert that a claim p is true is just to assert that p, and there is nothing more to a claim's being true than its being "assertible." A moral claim can be true even if it does not correctly describe moral *reality* in a way that moral *anti-realism* prohibits. (And, says the minimalist, what goes for "truth" goes for "facts," not surprisingly since these terms are generally taken to be interdefinable. True claims state facts, and statements of fact are true.) Although Timmons is a minimalist, he recognizes that there is a sense in which his brand of anti-realism rules out the existence of moral truths and moral facts. To acknowledge this he adopts a convention according to which the anti-realist denies that there are moral TRUTHS and moral FACTS and that moral judgments correctly describe the WORLD and the moral PROPERTIES it contains, even though there are moral truths, moral facts, and moral judgments that correctly describe the world

and its moral properties. I prefer to avoid this manner of speaking,[10] but I can use it to clarify my own terminology. By "truths" I mean TRUTHS, and by "facts," I mean FACTS, etc.

The organizing questions of this essay are these:

1. Is moral realism true? Is there any fact of the matter about what's morally right or wrong?

2. If moral realism is false, what should we do about it? What consequences does anti-realism have for our moral practices and our moral discourse in particular?

The rest of this chapter is devoted to clarifying these questions and identifying some possible answers.

---

[10] I prefer to avoid it because it suggests that "truths" and "facts," etc. are not ordinarily taken to be TRUTHS and FACTS, etc.

1.2 Two Meanings of "Moral"

Ask a garden-variety Westerner what it is to be a moral person and you might get

an answer like this: "A moral person is one who thinks about what's right and

what's wrong, tries to do the right thing, and usually manages to do so." You

also might get an answer like this: "A moral person is one who doesn't just think

about herself, but who thinks about the interests of others and tries to take the

interests of others into account in her actions." I suspect that most people would

say that these answers are very similar, if not two different ways of saying the

same thing. After all, one might say, thinking about what's right and what's

wrong and trying to do what's right is, for the most part, just a matter of thinking

about the interests of others and trying to take them into account in one's

actions.[11] Consider the Golden Rule: Do unto others as you would have them

do unto you. One might say that morality, at its core, is simply about giving

sufficient consideration to the interests of others in one's thought and action.

If moral realism is true—i.e. if there really is a fact of the matter about

what's right and what's wrong—and if our ordinary assumptions about the

general content of morality are correct, then there is no problem with saying that

the two answers given above are roughly equivalent. Doing the right thing is just

a matter of taking into account the interests of others (in the appropriate way),

---

[11] At least in Western culture. See Section 3.5.

and, likewise, taking into account the interests of others (in the appropriate way) is *de facto* doing the right thing. But what if moral realism is false? What if there really is no fact of the matter about what's right and what's wrong? In that case, the differences between these two accounts start to matter, especially for people who believe that there is no fact of the matter about what's right and what's wrong.

Consider the case of Brett. Brett is an anti-realist. He denies that there is any fact of the matter about what's right or wrong. But he's not a psychopath. In fact, he is, by all the usual measures, a caring and considerate person. As it happens, Brett has made a promise that has become unexpectedly costly to keep, and now he must decide whether or not to keep it. He considers the interests of the person to whom he made his promise and how that person will feel depending on whether or not he breaks it. He considers how he would feel if he were in that person's shoes. He considers his own interests and the effects that breaking or not breaking his promise will have on himself. He considers all of these things not as means of paying lip-service (or "mind-service") to accepted deliberative norms, but because he sincerely cares about them. After deliberating for what he takes to be a reasonable amount of time, he makes his decision, and it's a decision that most people would think reasonable. At no point does Brett aim to do "the right thing," as he does not believe that there is a "right thing" to be done. Instead he examines the situation in light of his values and in light of the facts as he understands them and then acts accordingly.

Suppose this sort of behavior is typical of Brett. Is he a moral person? He is according to the second account given above, as he is very much concerned with the interests of others and with making sure that his actions reflect those interests. Given what we know about Brett, the second account seems to get things right. Brett is, in the sense that matters to those who might know him, a moral person. But according to the first account given above, Brett is *not* a moral person. Brett has had no thoughts at all about right and wrong and has made no effort whatsoever to do what is right. This suggests that there is something wrong with the first account.

If moral realism is correct, i.e. if there is a fact of the matter about right and wrong, the problem is easily solved. Brett can be thinking about right and wrong even though he doesn't know it, even though he explicitly denies that that is what he is doing. Under these circumstances, Brett's case is similar to that of Lois Lane who is, in a very real sense, dreaming about Clark Kent when she dreams about Superman, even though she doesn't know it and even though she would balk at the suggestion that that is what she has been up to. Lois can dream about Clark without knowing it because there is, unbeknownst to her, an intimate connection between him and the heroic figure of whom she is so fond. Moreover, Lois' ability to dream about Clark in this way is made possible by the fact that she is a character in a fiction according to which Superman is *real*. We, however, can't dream about any real person by dreaming about Superman, nor can we dream about Superman by dreaming about some real person. (Christopher Reeve doesn't count.) The fact that Superman is a fictional

character rules this out, even for those of us who have been taken in by the fiction.

Thus moral realism can save the first account, allowing even an avowed anti-realist such as Brett to think about right and wrong without knowing it. However, if realism is false, if "right" and "wrong" are, like Superman, mythical creatures, then the first standard is once again in trouble, as Brett can no longer be considered a "moral person" given (a) his lack of interest in doing what is "right" and (b) the absence of any connection between Brett's thoughts and intentions on the one hand and some first order moral facts on the other. Nevertheless, I claim, Brett *is* a moral person in the sense that really matters. His unusual meta-ethical beliefs, mistaken or not, seem largely beside the point when it comes to evaluating the quality of his character.[12]

I draw two lessons from the above discussion. First, there are (at least) two different senses of "moral." These two senses tend to overlap under an assumption of moral realism, but they come apart when realism is called into

---

[12] Another illustration of the same point: In the documentary *When We Were Kings* (1997), the infamous fight promoter Don King is described by one of the interviewees as "amoral." It is unlikely that this interviewee was merely attributing to King an anti-realist meta-ethical view. Rather he meant that King is untrustworthy and ruthless, that he would do anything for a buck, etc. He calls King "amoral" because of King's lack of regard for the interests of others, a lack of regard that is compatible with both realist and anti-realist meta-ethical commitments on King's part.

question. Second, it is the second sense of moral, the one according to which

morality is a matter of having a sufficiently high level of consideration for the

interests of others, which captures what is essential to being a "moral person."[13]

These lessons can be resisted. I've used the example of Brett, a person

who has no interest in right and wrong but a deep concern for the interests of

others, to pull these two senses of "moral" apart. One might question my attempt

to separate these two senses, not on the grounds that moral realism is true *per*

se, but on the grounds that the story told above about Brett is incoherent. The

objection is that Brett couldn't maintain his commitment to considering the

interests of others were he to *really* believe that there is no fact of the matter

about right and wrong. In other words, being a decent person requires some sort

of commitment to moral realism.[14] I will consider this objection in greater detail in

Section 4.1.2. In the meantime, the psychology attributed to Brett is, I trust,

plausible enough, and I will work under the assumption that it is coherent and

that we are warranted in accepting the two lessons drawn above.

To give ourselves some terminology that reflects our distinction between

the two sense of "moral" outlined above we can define the following two terms.

moral$_1$: of or relating to the facts concerning right and wrong, etc.

---

[13] Again, this is more true in Western culture than in others. See Section 3.5.

[14] Point due to David Sussman.

This is the "moral" that is distinctive of moral realism. The "etc." indicates that other moral concepts besides that of rightness and wrongness fall under the moral₁ heading as well, provided that they are interpreted in a realist way. For example, one makes a moral₁ claim by calling some institution "unjust" or by saying that one has a "duty," "right," or "obligation" to behave in a certain way, provided, once again, that these terms are interpreted in a suitably realistic fashion.

Alternatively, we have:

moral₂: of or relating to serving (or refraining from undermining) the interests of others.[15]

(Note that neither of these definitions reflects a different, but perfectly ordinary use of "moral" by which it means "morally good" or "morally commendable." E.g. "You did the *moral* thing."[16])

---

[15] Morality is not exclusively limited to behavior that affects the interests of others. For example, some say that it would be wrong to do certain things, e.g. to commit suicide, even if you were the only being in existence, and in other cultures there is greater emphasis on moral violations do not have any obvious victim. (See Section 3.5.) For our purposes however, we can ignore these complications since the point here (that being "moral" does not necessarily involve being a moral realist) may be made in either case.

Ordinarily we do not distinguish between morality$_1$ and morality$_2$. (Hence the need for artificial subscripts. No pair of words is easily co-opted for the job.[17]) I take this as an indication of the pervasiveness of moral realism. The practice of trying to act rightly inevitably subsumes the practice of taking into consideration the interests of others. In other words, we, as realists, behave morally$_2$ simply by doing our best to behave morally$_1$.

---

[16] Note that when we considered whether or not Brett is a "moral person" we may have been using "moral" in a sense that merges this third sense with moral$_1$ and/or moral$_2$. If we say that Brett is a "moral person" we are saying that he engages in (or is capable of engaging in) moral$_1$ and/or moral$_2$ reflection, but we may be saying something further. We may be *commending* him for being that way or making a first order moral claim to the effect that being that way is morally$_1$ good.

[17] "Altruism" as a synonym for "morality$_2$" is on the right track, but not quite good enough because "altruism" refers primarily to active helping behavior whereas morality$_2$ is concerned with refraining from harming as much as with active helping. As far as I can see there is no ordinary term that is close to equivalent to *morality$_1$*.

1.3 Some Further Clarifications

Above we said that moral anti-realism is the thesis that there are no first order moral facts/truths, but, having not yet distinguished between morality$_1$ and morality$_2$, we did not say whether we were talking about putative moral$_1$ facts/truths or putative moral$_2$ facts/truths. It should now be clear that the first order claims in question are putative moral$_1$ facts/truths since they are claims concerning which things are right, wrong, and the like. In fact, we can think of moral realism as the thesis that morality$_1$ is for real, and moral anti-realism as the thesis that it is not.

We are interested in the consequences of anti-realism for our moral practices and discourse. What we mean by "moral practice" and "moral discourse" will depend on which meaning of "moral" we adopt. For now we can remain neutral by saying that *moral practices* are those behaviors performed by rational agents that are of moral concern without choosing between the two senses of "moral."

When it comes to what counts as a moral practice we are democratic. If enough people (perhaps one is enough) think that some behavior is a matter of moral concern then it *is* of moral concern and therefore is a moral practice in this very broad sense. (A consequence of this is that whether or not something is a

moral practice depends on the psychological states of the members of the relevant community.)

Note once again that, according to this usage, to say that something is a "moral practice" is not to commend it in any way, but merely to indicate that it is of moral concern. Thus, the practice of executing alleged witches is, in this sense, a moral practice.

We'll say that *moral discourse* is communication concerning one's moral beliefs and/or values with the aim of guiding moral practices. Moral discourse takes place using a *moral language* or a *moral vocabulary*, which is a part of the natural language in which the moral discourse takes place.

In what follows, where "moral" appears without a subscript the reader can assume that both interpretations are appropriate. The subscriptless "moral" will often appear when we are discussing our *present* moral discourse and practices since, as noted above, we don't ordinarily distinguish between the two senses of "moral." In a realist culture such as ours, the moral$_1$ discourse and practices coincide respectively with the moral$_2$ discourse and practices.

In what follows I will consider a number of views concerning the practical consequences of anti-realism. I divide the responses into two general categories: conservative and revisionist.

1.4 The Conservative Approach

According to proponents of the *conservative* approach, the truth of anti-realism provides us (or would provide us) with no general reason to revise our moral discourse and practices.

1.4.1. The Non-Cognitivists

During the 20th Century, many philosophers began to have doubts about moral realism. One response to these doubts was the emergence of *non-cognitivism*, according to which moral judgments are not descriptive, which is to say that they do not state facts and are not truth-apt. Moral judgments, say the non-cognitivists, may *appear* to be descriptive, fact-stating (when accurate), and truth-apt, but they are in fact expressions of approval/disapproval (Ayer, 1952), commands of some kind (Hare, 1952),[18] or both (Stevenson).

     The non-cognitivists took their enterprise to be one of *understanding* our moral practices and discourse *as they are*, and not one of *revising* them.[19] That's

---

[18] Hare, like Timmons (1999), takes something of a minimalist view of moral truth, and as a result may say that there are moral facts/truths. All the same, his view does not appear to be compatible with moral realism in the sense defined above.

[19] At least that's the standard interpretation. It seems to me, however, that Ayer's (1952) account is very error-theoretical in spirit. Witness the title of the relevant

not to say that these philosophers lacked first order moral agenda. Hare, for example, is a staunch defender of utilitarianism (1981). Nevertheless, these philosophers did not see the truth of non-cognitivism as providing any general reasons for changing the way we talk about moral issues or for changing our minds about any particular moral issues.[20] In this regard, contemporary non-cognitivists such as Alan Gibbard (1990) tend to take stances similar to those of their predecessors.

1.4.2. Mackie

Mackie's view combines the anti-realism traditionally associated with non-cognitivism and the cognitivism traditionally associated with realism. He agrees with moral realists that first order moral claims are intended to be descriptive, fact-stating, and truth-apt, while agreeing with non-cognitivists that moral claims never in fact describe the world, state facts, or qualify as truths. We think that moral claims are sometimes true, but we are mistaken, says Mackie; our moral thought and language presuppose the existence of objective values which, as it happens, do not exist.

chapter of *Language, Truth, and Logic*: "Critique of Ethics and Theology." One gets the feeling that Ayer is not merely interested in explicating the nature of moral talk, but—to some extent, at least—in debunking it as well.

[20] Once again, Ayer (1952) may have been an exception.

Mackie clearly advocates a change in the way we *understand* our moral discourse and other practices. We should give up our belief in objective values and bear this in mind when we use moral vocabulary and participate in moral discourse. The extent to which Mackie advocates actually *changing* our moral discourse and practices is unclear. The consensus in the literature seems to be that Mackie does not advocate any such change. Rather, his interpreters tend to agree that Mackie would have us talk and act in the moral domain, more or less, as we do. J. A. Burgess (1998, pg. 535) offers the following reading.

> [Mackie] seems to have thought it fairly uncontentious that the acceptance of an error theory for moral language ought to incline us to change neither the content nor the form of our moral practices nor our attitude toward those practices.

Burgess doesn't define what he means by "practices" or "the content… of our moral practices." If the definition that I've given above is on target, then it's clear that Mackie does advocate some change in "the content of our moral practices." For example, he advocates infanticide for babies born with sufficiently severe, permanent defects (1977, pg. 199). One could argue, however, that insofar as Mackie does advocate changes in "the content of our moral practices" these changes are motivated by concerns that are largely independent of his meta-ethical views or, somewhat more mildly, that these first order commitments do not follow in any straightforward way from his meta-ethical commitments.

Blackburn has the following to say concerning the practical upshot of Mackie's view.

> Mackie did not draw the consequences one might have expected from this position…. [One might have expected him to say that] our old, infected moral concepts or ways of thought should be replaced by ones which serve our legitimate needs, but avoid the mistake. Yet Mackie does not say what such a way of thought would look like, and how it would differ in order to show its innocence of the old error. On the contrary, in the second part of the book, he is quite happy to go on to express a large number of straightforward moral views… All these are expressed in the old, supposedly infected vocabulary (1993a, pp. 149-150).

In the following passage Mackie does seem to describe—if only in very general terms—how a revamped moral discourse could show its innocence of the old error.

> My hope is that concrete moral issues can be argued out without appeal to any mythical objective values or requirements or obligations or transcendental necessities… (1977, pg. 199).

Perhaps he means by this nothing more than that an improved moral discourse will make no reference to the Will of God and the like, but that seems unlikely

given that Mackie's targets include atheist moral philosophers. One has to wonder exactly what sorts of appeals he'd like to eliminate. The list includes "mythical… requirements or obligations," but, in the course of discussing whether or not his normative view could properly be called "utilitarian," he says the following:

> [My approach would be] not just a rule-utilitarianism but a rule-right-duty-disposition utilitarianism (pg. 200).

It sounds as if Mackie's approach allows, if not encourages, appeals to "rights" and "duties" in spite of his rejection of appeals to "mythical… requirements or obligations." Given the conceptual connections between "rights" and "duties" on the one hand and "requirements" and "obligations" on the other,[21] it appears that the modifier "mythical" is crucial to his maintaining consistency. But which "rights" and "duties" are supposed to be mythical? From reading Part I of Mackie's book one easily gets the impression that they all are. If Mackie thinks that rights and duties divide into the mythical and the non-mythical, then it's rather odd that he neither mentions this explicitly nor explains how to tell the two apart. If none is mythical, then why is he worried about us appealing to the mythical ones? If they're all mythical, then how could we appeal to any rights and duties without appealing to mythical ones?

---

[21] See Kanger and Kanger (1966).

Once again, whether or not Mackie advocates substantial change in our moral

practices and discourse is unclear. He clearly has an agenda as a normative

ethicist, and the advancement of such an agenda would certainly change our

moral practices in the broad sense of "moral practices" defined above. Whether

or not Mackie would have our moral practices remain the same in some more

narrow sense of "moral practices" is an open question, and Burgess and

Blackburn suggest that he would. What is clear is that Mackie fails even to

gesture toward an alternative moral vocabulary or set of moral concepts to

replace the error-infected ones we have. Given that, it seems safe to say, as his

interpreters suggest, that Mackie advocates substantial change in neither the

general style of our moral discourse nor the vocabulary in which it is conducted.

1.4.3. Blackburn's Quasi-Realism

Once again, Blackburn calls himself an anti-realist because he believes that

moral judgments "start theoretical life" (1993a, pg. 168) as expressions of non-

cognitive attitudes rather than as descriptions of features of the world. He calls

himself a quasi-realist because he believes that moral judgments can "earn the

right" to be expressed using the propositional and apparently descriptive/realist

language that we ordinarily use to express moral judgments. He advocates

substantial change in neither our moral practices nor our moral discourse, though

he does favor a more sophisticated, projectivist anti-realist understanding of what

we are doing when we moralize. According to Blackburn, predicates such as "is right" and "is wrong," though importantly different from ordinary non-moral predicates such as "is over six feet tall," are sufficiently well-behaved to be retained and to play the action-guiding role that they typically play.

The comparison between Blackburn and Mackie is illuminating. Mackie believes that the truth of anti-realism undermines ordinary moral thought. As noted above, he is not entirely clear about what he takes the consequences of anti-realism to be for our moral practices and discourse, though the consensus among interpreters is, once again, that Mackie takes the consequences to be minimal. Blackburn is one such interpreter, and he views Mackie's lack of interest in revising our moral practices/discourse as casting doubt on Mackie's claim that ordinary moral thought and language are infected with error. While Mackie believes that the truth of anti-realism undermines ordinary moral thought (and perhaps to some extent ordinary moral practices/discourse), Blackburn views the acceptance of anti-realism as a deepening of our understanding of our moral thought, practices, and discourse and not as grounds for revising them.

1.4.4. Timmons' Ethical Contextualism

In his recent book *Morality Without Foundations* (1999) Mark Timmons defends a view that he calls *ethical contextualism*. Timmons' view is strikingly similar to Blackburn's. Both are anti-realists with ontological commitments similar, if not

identical, to Mackie's, and both deny, contra Mackie, that there is an error

embodied in ordinary moral thought and language. Both take a liberal approach

to truth, denying that true claims must (as a general rule) accurately describe a

mind-independent world. Instead, they argue that there is a legitimate notion of

truth that can apply to moral judgments within their anti-realist frameworks,

frameworks which have no room for mind-independent moral facts. For

Timmons, the truth of a moral judgment is a matter of its being "correctly

assertible," where a claim's correct assertibility depends not only on the mind-

independent world but on the semantic norms that are invoked by the context in

which it is made.[22] Like Blackburn, Timmons attempts to ward off the specter of

relativism (relativity to semantic context in Timmons' case) by appeal to a

distinction between an "engaged" and "detached" moral standpoint, the former

being the standpoint from which one can legitimately say that moral judgments

are categorical and not relative.

---

[22] Of course, even hard-line correspondence theorists recognize that the truth-

values of sentences depend on semantic context. For them, the truth value of

the sentence "I am seated" will depend on the semantic context in which it is

uttered. Timmons, however, maintains that moral truths are dependent in their

semantic contexts in a much deeper way.

<u>1.4.5. Moral Fictionalism</u>

Moral fictionalism is best understood as two distinct and complementary views. *Descriptive* fictionalism is error theory. According to this view, ordinary moral claims are all false, like those made in a fiction, but unbeknownst to their authors.[23] Mackie is a fictionalist in this sense, though he does not use this term.

Proponents of *prescriptive* fictionalism are descriptive fictionalists who nevertheless advocate speaking and acting as if some form of moral realism were true, presumably because maintaining the fiction has, on balance, desirable effects. What one might call a "democratic" prescriptive fictionalist advocates a widespread, self-conscious pretense that moral realism is true, while an "esoteric" prescriptive fictionalist advocates a system by which a select group of individuals self-consciously indulge in the realist fiction while allowing or even encouraging others to believe it.

Mackie does not describe himself as a prescriptive fictionalist, but Blackburn (1993a, pg. 150) and Hinckfuss (1987, pg. v) appear to suggest that he is. The most developed account of fictionalism appears in Richard Joyce's *The Myth of Morality* (2001). Joyce defends his version of error theory

---

[23] "Descriptive fictionalism" could also designate the rather implausible view that some or most people are already error theorists who have adopted prescriptive fictionalism. (See below.) Likewise, it could designate the view that descriptive fictionalism (in the sense just mentioned) describes the unconfused state that most resembles our present state of meta-ethical confusion.

(descriptive fictionalism) and then argues that maintaining moral$_1$ discourse and thought as a fiction may be advisable given our interests (prescriptive fictionalism). He argues that self-consciously maintained fictions can, and sometimes do, play an important role in a culture's practices and that avowed error theorists who think and speak as if moral realism were true stand to gain important benefits, chief among them an effective device for combating weakness of the will in moral$_2$ contexts. We will consider Joyce's views in greater detail in Chapter 4.

<u>1.4.6. Others</u>

There are a number of philosophers, writing both before and since Mackie, whose meta-ethical views are anti-realist, or whose views share certain affinities with anti-realism, even though they do not, and perhaps would not, describe themselves as "anti-realists." John Rawls (1971, 1980), for example, is a *constructivist* who denies the existence of moral facts/truths that are independent of our beliefs, reactions, or attitudes. (See Section 4.7.1.) Anscombe (1958), in a somewhat Mackiean vein, charges that "modern moral philosophy" is flawed in that its central concepts are rendered unintelligible by our present-day tendency to reject divine law. Foot (1972) defends the view (a view she has since rejected (1994)) that morality may be best understood as a system of "hypothetical imperatives," commands the authority of which depends on the values and commitments of those who might obey them. Others such as Williams (1985a)

and Warnock (1971) may fall in this category as well. What matters for our purposes is that none of these philosophers proposes any substantial revisions in the general form of our moral discourse and practices.

1.5 The Revisionist Approach

In contrast to the conservative thinkers described above, *revisionist* meta-ethicists believe that the truth of anti-realism does, or would, provide us with general reasons for revising our moral discourse and practices, at least given certain background assumptions.

1.5.1. Hinckfuss' Moral "Nihilism"

In a discussion paper entitled "The Moral Society: Its Structure and Effects," Ian Hinckfuss (1987) defends "moral nihilism." He takes nihilism to be the view that there are no moral obligations (pg. v).

Hinckfuss favors an "amoral" society, one in which the inhabitants do not believe in the existence of moral obligations and, naturally, do not strive to meet any. The practical upshot of this position depends crucially on what one means by "moral."

On the one hand, one might take "moral" to encompass both the $moral_1$ and the $moral_2$. Under this interpretation, an "amoral" society is one whose

members (a) believe that there is no fact of the matter about right and wrong and *a fortiori* make no attempt to behave in accordance with such facts and (b) place no special value on practices embodying consideration for the interests of others. On the other hand, one might take "moral" to refer more narrowly to matters moral$_1$. Under this interpretation, an "amoral" society would, as before, shun all talk of and concern with "right" and "wrong" etc., but might nevertheless place great value on practices embodying consideration of the interests of others, i.e. on morality$_2$. Its members would be like Brett and, from what I'm told, Hinckfuss himself. (I take it for granted that "moral" must refer, at least in part, to matters moral$_1$ and not exclusively to matters moral$_2$.)

Although Hinckfuss' language frequently suggests the first and more radical interpretation, I'm confident that what Hinckfuss really opposes is morality$_1$, not morality$_2$. I should emphasize, however, that Hinckfuss does not explicitly distinguish between the two senses of "moral"[24] and, as one might

---

[24] Hinckfuss (pp. 14-15) does make a distinction between "immorality" and "amorality" which, in my terms, is the distinction between "immorality$_1$" and "amorality$_1$." While this distinction captures an important difference between two possible ways of relating to morality$_1$, it does not capture the difference between morality$_1$ and morality$_2$.

As Hinckfuss uses the term—and it's a perfectly ordinary use—"immorality" depends on the truth of moral realism. To do something "immoral" is to contravene the moral facts. But what I call "amorality$_2$" or "immorality$_2$" does not depend on moral realism for its application. One who places no special value

expect, he does not have much to say about what sort of moral$_2$ life he envisions for his amoral$_1$ society. He says next to nothing about the moral$_2$ practices and discourse it would support. We can infer from his amoralism$_1$ that members of Hinckfussian amoral society will neither use characteristically moral$_1$ language nor appeal to moral$_1$ considerations. Hinckfuss does make one positive recommendation for the shaping of moral$_2$ life. He advocates "rational resolution of conflicts" whereby disputants first attempt to clarify the facts relevant to their dispute and sort out any conceptual misunderstandings between them and then, if necessary, attempt to devise mutually acceptable ways of resolving their conflict.[25]

---

on the interests of others is amoral$_2$ or perhaps immoral$_2$, and such people can exist even if there are no moral facts.

The difference between these two distinctions will prove to be important. If moral realism is false, then once one does away with what Hinckfuss calls "morality" (i.e. morality$_1$) there is no practical need to think or talk about the extent to which we want to be (im)moral$_1$ because it's impossible to be (im)moral$_1$ if moral realism is false. But, even if we reject moral realism and morality$_1$, we still are left with a question concerning how we will think and talk about our moral$_2$ concerns—a very important question in my opinion and those of many other philosophers, e.g. Blackburn (1984, 1993a).

[25] For an earlier view similar to Hinckfuss', see Stirner (1973).

1.5.2. John P. Burgess' "Anethicism"

In an unpublished paper entitled "Against Ethics," John P. Burgess (1979) defends a view he calls "Anethicism." According to Burgess "Anethicism is to ethics as atheism is to theology," (Pg. 1). Like Mackie and Hinckfuss, J. P. Burgess claims that there are no objective values, in spite of the fact that ordinary moral thought and language presuppose that there are. Unlike Mackie, and like Hinckfuss, J. P. Burgess thinks that the truth of anti-realism gives us reason to change the nature of our moral discourse and our moral practices. While J. P. Burgess is happy to talk about personal ideals and group mores, he thinks that we should drop all talk of "right" and "wrong" and any other vocabulary that furthers the illusion that there are objective values. In short, J. P. Burgess suggests that we eliminate morality$_1$. The extent to which he favors the retention of morality$_2$ depends on the content of the personal ideals and group mores he favors, a topic on which he does not comment.

1.5.3. Garner's "Compassionate Amoralism"

Richard Garner defends a version of revisionism in his book *Beyond Morality* (1994). Garner, like Mackie, Hinckfuss, and John P. Burgess, believes that ordinary moral thought and language are deeply flawed in that they mistakenly presuppose the existence of objective values. Like Hinckfuss and J. P. Burgess,

he argues that we should do away with all moral talk and all appeal to moral considerations. He suggests that we replace our interest in morality with "compassion, a desire to know what is going on, and a disposition to be non-duplicitous," (Pg. 3).

Garner calls himself an "amoralist" and his book "Beyond Morality." These facts, taken in isolation, might suggest that Garner is opposed to morality in its entirety, i.e. to both morality$_1$ and morality$_2$, but this is not so. Garner favors only amoralism$_1$, not amoralism$_2$. His ideal society is one in which its citizens care deeply about each others' interests. He believes that adopting the "amoral" position prevents one from being distracted by the confusing and endless squabbles over moral principles and moral language and allows one to attend to the relevant facts in making one's decisions.

1.5.4. Radical Nihilism

I use the term *radical nihilism* to cover a family of views united in their rejection of both morality$_1$ and morality$_2$. It is this family of views that ordinarily comes to mind when one hears the word "nihilism."[26] (Hinckfuss calls himself a "nihilist,"

---

[26] John P. Burgess (Pg. 17) agrees. "There are some people, called *sociopaths*, who lack all conscience and community feeling. They are delinquents and criminals and are generally held in low esteem by law-abiding citizens. Yet one school of moral philosophers think we should all be like them. These are the

but is not, as I read him, a "radical nihilist," and thus not a "nihilist" as the term is typically used.) Radical nihilists not only reject moral realism (i.e. morality$_1$) but are also either indifferent to or hostile toward morality$_2$. The indifferent radical nihilist says, "Do whatever you want. If you care about the interests of others, fine. If you don't, that's fine, too." The hostile radical nihilist goes even further, proclaiming that morality$_2$ is for the weak and that any consideration for the interests of others is pure foolishness.

As far as I know, no noteworthy philosophers within the Anglo-American tradition endorse radical nihilism in either form. (Nietzsche, depending on how one interprets his views, may be considered a radical nihilist, and as such would be its most prominent proponent.) Radical nihilism warrants some attention, however, not because of its appeal, but rather because a number of philosophers suggest that it is an unwitting, but nonetheless inevitable, consequence of anti-realism and/or error theory. Blackburn writes:

> It is, I think, undeniable that moral anti-realism is often seen as a dangerous doctrine, a more or less surreptitious denial of the importance of ethics… It is thought to consort with lack of real seriousness, just as relativism is felt to undermine any real commitment (1993a, pg. 208).

---

nihilists, whose motto is 'Fais ce que voudras!'" Note also that in the Coen Brothers' film *The Big Lebowski* (1998) Bunny Lebowski assures The Dude that her nearby male friend will not object to The Dude's blowing on her toes: "Uli doesn't care about anything. He's a nihilist."

Elsewhere he writes:

> A recent influential book even believes that an emotivist should approve of
>
> manipulating people, bullying and lying and brainwashing as we please,
>
> rather than respecting their independence… books and sermons alike
>
> pronounce that the projectivist [anti-realist] should, if consistent, end up
>
> with the morals of a French gangster (1984, pg. 197).

The "influential book" to which Blackburn refers is MacIntyre's *After Virtue* (1981).
MacIntyre's comments concern emotivism, but, as MacIntyre explains, the
feature of emotivism that makes these comments relevant is its failure to
recognize an impersonal source of moral authority. Thus, the following
comments apply equally well to all forms of anti-realism.

> What is the key to the social content of emotivism? It is the fact that
>
> emotivism entails the obliteration of any genuine distinction between
>
> manipulative and non-manipulative social relations. Consider the contrast
>
> between, for example, Kantian ethics and emotivism on this point. For
>
> Kant—and a parallel point could be made about many earlier moral
>
> philosophers—the difference between a human relationship uninformed by
>
> morality and one so informed is precisely the difference between one in
>
> which each person treats the other primarily as a means to his or her ends

and one in which each treats the other as an end. To treat someone else as an end is to offer them what I take to be good reasons for acting in one way rather than another, but to leave it to them to evaluate those reasons. It is to be unwilling to influence another except by reasons which that other he or she judges to be good. It is to appeal to *impersonal criteria* of the validity of which each rational agent must be his or her own judge. By contrast, to treat someone else as a means is to seek to make him or her an instrument of my purposes by adducing whatever influences or considerations will in fact be effective on this or that occasion (Pp. 23-24). [Emphasis mine]

MacIntyre's suggestion is that if you reject all *impersonal criteria* for evaluative judgment, that is, if you reject morality$_1$ and the objective values it requires, then you have no choice but to *use* people, to treat them as mere instruments for the satisfaction of your own desires. That is, you must reject *morality$_2$*. This claim should arouse suspicion in light of the previous discussion. The case of Brett suggests that one can reject moral realism without becoming a *rat-bastard*. I noted above that some would question the coherence of the psychology attributed to Brett and, more generally, the coherence of distinguishing morality$_1$ from morality$_2$. We will revisit this issue in Section 4.1.2. The point for now is that some philosophers believe that radical nihilism is an inevitable consequence of anti-realism.

Further evidence for the perceived marriage between anti-realism and radical nihilism lies in the fact that philosophers often view Mackie's error theory as a theory of "last resort" (McNaughton, 1988, pg. 98). Take for example Michael Smith in *The Moral Problem*. Smith sees the central problem of meta-ethics as one of finding a way to hold on to the following three claims: 1. Moral judgments purport to be objective. 2. Moral judgments are intrinsically motivating. 3. One is motivated to act if and only if one has an appropriate desire and an appropriate means-end belief, where this belief and this desire are distinct existences ("Humean psychology"). Smith's book is devoted to showing how one might hold on to all three claims and thus solve the moral problem, but Smith acknowledges from the outset that there is a well-known view that straightforwardly solves the problem as he describes it, namely Mackie's error theory. Mackie maintains that moral judgments purport to be objective and that they are intrinsically motivating, and his view is compatible with Humean psychology. Smith doesn't attempt to find an inconsistency in Mackie's view, nor does he attempt to weigh the costs and benefits of his view vis-à-vis Mackie's. He simply takes it for granted—as do most of his readers, I suspect—that Mackie's view is to be embraced only as a last resort.[27] Smith's framing of the problem suggests that to embrace Mackie's error theory is not so much to solve

---

[27] Kieran Setiya (personal communication) makes a compelling case that Mackie's view does not solve Smith's moral problem. Whether or not that's so, the point here depends only on Smith's believing that Mackie's view is successful in this regard.

the moral problem as it is to admit that it's unsolvable. In a similar vein, Darwall,

Gibbard, and Railton in their history of recent meta-ethics (1992) outline what

they take to be the major meta-ethical contenders at the end of the Twentieth

Century. Mackie's error theory is mentioned only briefly, suggesting that his

view, though a well-known fixture of Twentieth Century meta-ethics, is not a view

with much of a future.

Why is Mackie's error theory so readily dismissed? I suspect it has

something to do with the specter of radical nihilism. Indeed, Smith (1994) ends

his initial exposition of the moral problem with the following words: "Moral

nihilism… looms." But what, exactly, looms? A counterintuitive metaphysical

theory? What battle-worn philosopher hasn't gone in for one of *those*? The fact

that Mackie's view is so widely regarded as a theory of last resort suggests that it

is regarded not merely as implausible, but as downright *unsettling*. Smith's use

of the word "nihilism,"—as opposed to, say, "cognitivist irrealism"—is, I trust, no

accident.   If I'm right when I say that "nihilism" is, *pace* Hinckfuss, properly read

as "radical nihilism," then what supposedly looms is unsettling indeed. Smith's

account suggests that unless the moral problem can be solved by some other

means, then, insofar as we are honest with ourselves, human life as we know it

must come to a cold and bitter end. In other words, our ultimate failure to solve

the moral problem without resort to an error theory would spell doom not only for

$morality_1$ but for *$morality_2$* as well.[28]

---

[28] The distortion that occurs when one assimilates anti-realism to nihilism is

brought out by considering a similar assimilation in the case of aesthetics (J. P.

43

Whether or not Smith and others are right about the connection between anti-realism and radical nihilism will most likely depend on the now familiar issue concerning the separability of morality$_1$ and morality$_2$. If the two should prove inseparable then the anti-realist's rejection of the former will amount to a rejection of the latter as well, and hence an implicit commitment to radical nihilism.

Burgess, 1979). Many adults come to believe that "beauty is in the eye of the beholder," that what is or is not beautiful is ultimately subjective, and that after a certain point there is no point in arguing about what is or is not beautiful since there is really no fact of the matter concerning such things. Shall we dub those who hold this rather commonplace view "aesthetic nihilists?" To do so is to suggest that such people have no feel for or interest in beauty, that they would just as soon trade filet mignon for dog food and a cottage overlooking the sea for a shack next to a garbage dump. But, of course, there's a big difference between taking a thoroughgoing subjectivist view of aesthetics and wanting to live like a swine. Likewise, there's a big difference between being a moral anti-realist and being the sort of person that comes to mind when one hears the word "nihilist." Once again, the anti-realist can reject morality$_1$ without abandoning morality$_2$.

1.5.5. John A. Burgess' Framework for the Evaluation of Error Theories

John A. Burgess (1998) (not to be confused with John P. Burgess) has developed a framework within which error theories, and Mackie's error theory in particular, may be evaluated. More specifically, he is interested in the practical consequences of Mackie's error theory. He wants to know whether or not our moral practices and discourse can be preserved if something like Mackie's error theory is correct. Although he does not ultimately take a stand on this issue, his work deserves some attention since he is one of the few philosophers to address this issue explicitly.

J. A. Burgess believes that Mackie's critics often move too hastily to the conclusion that Mackie cannot (in good faith) advocate the preservation of our moral practices and discourse. He also thinks that Mackie is too complacent in his uncritical willingness to maintain our moral practices and discourse in light of his meta-ethical commitments. To sort the matter out, J. A. Burgess suggests that we attend to the "intrinsic" and "extrinsic" merits of this error-ridden discourse.

The first intrinsic merit that an error-ridden discourse might have is its being "localized," isolated from the rest of our discourses and practices in such a way that it doesn't infect them with the error. The second intrinsic merit such a discourse might have is its sentences' being such that they are "as good as… [or] better than the truth" for those who knowingly indulge in their use. For example, engineers may find that sentences in the language of Newtonian physics serve

their purposes as well or better than their more recently devised replacements. According to J. A. Burgess, the only extrinsic merit an error-infected discourse might have is its being "pragmatically or socially useful."

J. A. Burgess concludes that whether or not Mackie's error theory would, if true, require some revision of our moral discourse and practices depends on (a) the extent to which the error embodied in ordinary moral thought/vocabulary is localized and (b) the intrinsic and extrinsic (de)merits of carrying on with our current moral practices/discourse. In Part II I will assess the advantages and disadvantages of maintaining ordinary moral thought and discourse.

1.6 In Defense of Revisionism

My aim is to defend a version of revisionism in the spirit of Garner, Hinckfuss, and J. P. Burgess. More specifically, I advocate the simultaneous retention of morality$_2$ and rejection of morality$_1$. In terms of practice, I propose that we eliminate all talk of "right," "wrong," "rights," "duties," "obligations," etc. except in contexts in which it is clear that the application of such terms does not presuppose the truth of moral realism. I suggest instead that we replace our current moral talk with increased talk about morally$_2$ relevant facts and a revamped vocabulary for moral$_2$ evaluation that is transparently subjective. I advocate this radical change in the way we think and talk about morals for two main reasons.

First, I believe that giving up realist moral talk will allow our language to accurately reflect the nature of morality. Moral anti-realism is true, as I will argue in Chapter 2, and our moral language might as well reflect this truth. Second and more importantly, I believe that our giving up realist talk and realist thought will make the world a better place. Better in what sense? Objectively better? No. Subjectively better—better according to me and my values and, I suspect, better according to you and yours. I say that getting rid of realist moral thought and talk, i.e. getting rid of morality$_1$, will make the world a happier, more peaceful, and more compassionate place. Morality affords us incalculable benefits, but I maintain that the benefits of morality are, for the most part, the benefits of morality$_2$ and that we can enjoy them without the damaging effects (and false metaphysics) of morality$_1$. We can maintain the moral practices that improve human life while getting rid of those that don't, all without the help of morality$_1$.

Whether or not this is true is an empirical matter. Why, then, am I defending this claim in a philosophical essay? To begin, I, like many philosophers, reject any sort of sharp distinction between philosophical and empirical investigation. Philosophy may be "*a priori*" in the sense of "from the armchair" but it is not, or need not be, *a priori* in the Kantian sense. That is, philosophical theories need not be constructed and defended independently of contingent matters of fact. Second, before a question can be investigated empirically it must be raised and shown to be worthy of empirical investigation, and these tasks can be accomplished using traditional philosophical methods. Along these lines, my more modest aim is to raise a set of empirical questions

concerning the likely effects of life with morality$_2$ but without morality$_1$ and to argue that these questions are worth investigating. My more ambitious aim is to argue that jettisoning morality$_1$ while retaining morality$_2$ is an excellent idea.

The revisionist position, though far from popular, has been defended before. Once again Garner, Hinckfuss, and J. P. Burgess have argued, in not so many words, that we should do away with morality$_1$ while retaining morality$_2$. My disagreements with these three philosophers are, with one exception, relatively minor. My primary goal with respect to them is to complement and expand. I'll now say a few words about where this essay is headed and how the terrain it covers is different from that covered by these three philosophers.

I have no substantial disagreement with J. P. Burgess. My only complaint is that J. P. Burgess calls his view "anethicism" and claims to be "against ethics." I think this wildly misrepresents our view with respect to morality$_2$. (That is, unless, Burgess is, in fact, indifferent or hostile to morality$_2$, in which case his view and mine are substantially different.) Putting aside this small difference, I regard J. P. Burgess' work as a point of departure.[29] Burgess says little or nothing about each of the following questions concerning how a society that rejects morality$_1$ but accepts and values morality$_2$ will conduct itself.

How will such a society carry out its discussions of moral$_2$ issues?

---

[29] His comments concerning revisionism are limited to a few paragraphs at the end of his unpublished manuscript.

Is there nothing to discuss?

If there is, what sorts of discussions will take place, and with what

vocabulary?

How will its inhabitants resolve their conflicts?

How will its inhabitants raise their children?

Will they receive explicit moral$_2$ education? A temporary moral$_1$

education?

How should one who would like to transform our society into an amoral$_1$

one speak and conduct herself in the present circumstances?

I will address these questions among others in Part II, building on the discussions

in Chapters 2 and 3.

My disagreements with Hinckfuss are somewhat more substantial. First,

as with J. P. Burgess, I take issue with some of Hinckfuss' choice of words.

Hinckfuss calls himself a "nihilist" and advocates living in an "amoral" society.

"Nihilism" is usually synonymous with what I've called "radical nihilism," and,

thus, to call oneself a "nihilist" and proclaim oneself an enemy of "morality" is to

invite the contempt of anyone who recognizes the obvious importance of

morality$_2$.

I believe that Hinckfuss does distinguish implicitly between morality$_1$ and

morality$_2$, but his separation of the two is incomplete, as suggested by his

account of the undesirable effects of "morality." Hinckfuss maintains that a moral

society inevitably generates "elitism, authoritarianism, guilt complexes, ego

competitions, economic inequality, and war," (Pg. v). I maintain that these problems are generally exacerbated by morality$_1$, but often mitigated by morality$_2$, and therefore it is important to separate morality's components in assessing the damage done by "morality."[30] One of my aims vis-à-vis Hinckfuss is to give a more measured account of the advantages and disadvantages of revisionism.

Like Hinckfuss, I favor the demise of morality$_1$ and believe that the benefits of abandoning it far outweigh the costs. Still, I think Hinckfuss' treatment of this issue does not do justice to the potential benefits of moral$_1$ practices and discourse and the potential costs of abandoning them. One might consider, for example, whether or not one can effectively raise children without the use of moral$_1$ language. Likewise one might wonder whether or not leaders of worthy causes can effectively rally support without the rhetoric of morality$_1$. (For example, could the substantive ends of the human rights movement be served without talk of "rights" and the like?) Finally, one might wonder along with Richard Joyce (2001) whether moral$_1$ language and thought can be an effective device for combating weakness of the will.

There are other questions that Hinckfuss largely ignores. For example, the members of his "amoral" society will need some way of talking about evaluative matters concerning behavior. That is, they need some anti-realist replacement for moral$_1$ talk. Hinckfuss, however, proposes no such replacement

---

[30] In Chapter 4 I argue based on psychological considerations that getting rid of morality$_1$, while likely to have many beneficial effects, is not a recipe for eliminating all the ills of human nature.

talk. (He thinks that the influence of "morality" is so pernicious that the question of what to put in its place is one that we can comfortably put aside.) Moral$_1$ discourse, even if error-ridden, performs an important function in our society. I suggest, contra Hinckfuss, that the revisionist does need to offer either (a) some sort of replacement for moral$_1$ discourse or (b) some explanation of how we can do without one. I will attempt to provide a little of each. Some aspects of moral$_1$ discourse we can replace; others we can do without.

To be fair, Hinckfuss does go some of the way toward providing a replacement discourse, specifically with respect to conflict resolution. However, his program for resolving conflicts (a) does not (and is not intended to) offer guidance to those who would discuss moral$_2$ matters in the absence of conflict and (b) offers no guidance for discussion in cases of conflict once it's been determined that the conflict in question is not simply a by-product of factual error or conceptual misunderstanding, which, I maintain, is typically case. One of my aims is to suggest ways in which members of an amoral$_1$ society might discuss matters of morality$_2$ in general (i.e. not just in cases of conflict) and how they might deal with conflicts that arise from underlying conflicts of value.

Hinckfuss argues that morality$_1$ exacerbates conflict, a point that I take to be extremely important. Hinckfuss cites one sociologist's work (Trainer, 1982) in support of this point. One of my primary aims in this essay is to expand on Hinckfuss' work by drawing further empirical support for this claim.

With Garner, as with J. P. Burgess, I have no substantial disagreements. As with J. P. Burgess and Hinckfuss, I take issue with Garner's choice of words.

Garner calls his view "compassionate amoralism." In terms of PR this is much better than "nihilism" and "anethicism," but "amoralism" nevertheless invites misinterpretation, even if "compassionate" undoes most of the damage.[31] Similar remarks apply to the title of his book *Beyond Morality*.[32] All the same, it's clear enough that Garner embraces morality$_2$ in his rejection of morality$_1$.

<hr />

[31] George W. Bush may have since ruined the term "compassionate."

[32] I claim that Burgess, Hinckfuss, and Garner misrepresent the view we share though similar poor choices of words. That's three against one, and I would be a bit myopic were I not to consider the possibility that it is *I* whose words misrepresent the position we share. Nevertheless, I stick to my guns. First, it's not really three against one. These three do not present a united linguistic front, and there are others who speak as I do, mostly realists. Burgess (pg. 17), for example, also identifies nihilism with radical nihilism. Moral realists such as Michael Smith reject "nihilism" out of hand and without justification. This makes sense if one takes "nihilism" to entail a rejection of morality$_2$ as well as morality$_1$. But if "nihilism" means what Smith, Burgess, and I say it means, why would a friend of morality$_2$ such as Hinckfuss call himself a "nihilist?" Likewise, why does Burgess say that he's "against ethics," and why does Garner favor "amoralism?" I think the answer has to do with lack of terminological alternatives. In general, when one needs to draw a distinction, one's first inclination is to use existing words in order to make it, even if meanings must be stretched a bit. It seems to me, however, that, due to the pervasiveness of moral realism, there are no good words in English for distinguishing between morality$_1$ and morality$_2$, and hence no

As noted above, I will attempt to support my claim that the world would be better off without morality$_1$ by appeal to empirical considerations, primarily from psychology. In Chapter 4 I will introduce my revisionist proposal for how to think and talk about moral$_2$ issues and argue on pragmatic grounds against the views of conservative anti-realists. In Chapter 5 I will address some of the practical challenges facing revisionism as a program of social reform and defend utilitarianism, not as a normative theory of ethics, but as a guideline for practical judgment in the public domain.

The defense of revisionism offered in this essay depends on the highly controversial assumption that anti-realism is true. Alas, I've no knock-down argument against moral realism, but I think it important nonetheless that the reader be exposed to the sorts of considerations that have led myself and others to reject it. To that end, I make my case for anti-realism in the next chapter. In Chapter 3 I will complement Chapter 2 by developing a projectivist account of moral psychology and with it an explanation for why moral realism is so appealing in spite of the fact that it's false.

---

good words for describing a view that embraces morality$_2$ while rejecting morality$_1$. In the name of clarity, I opt for the clunky, artificial subscripts, but I can see why one would be inclined to do otherwise.

Chapter 2

Against Moral Realism

In "The Many Moral Realisms" Geoffrey Sayre-McCord (1988) writes that "the plausibility of anti-realism… depends largely on preserving our normal ways of speaking even while challenging the natural (though perhaps naïve) realistic interpretation of what is being said." I couldn't disagree more. To my mind, the plausibility of moral anti-realism lies in the *implausibility of moral realism*, and our normal ways of speaking be damned if they no longer make sense after moral realism has been rejected.

In this chapter I will explain why I find moral realism so implausible. I should emphasize that what follows is not a thorough defense of anti-realism. Rather, my aim is to advance, in broad outline, what I take to be the most compelling reasons for rejecting moral realism and thus to help the reader understand what would lead a reasonable person to reach the conclusions I've reached, conclusions which are very much at odds with common sense.

2.1 Some Preliminaries

2.1.1 Two Sources of Moral Skepticism

Let us call the rejection of moral realism "moral skepticism."[1] Moral skepticism

has at least two distinct sources, and if we are to take moral skepticism seriously

we must distinguish between them.

Some people are moral skeptics because they are dissatisfied with

conventional moral life: "Who are *they* to tell *me* how to live? What makes them

so sure that their values are anything more than *their values*?" Such a skeptic

comes to doubt that there are moral truths because he is dissatisfied with what

those around him take the moral truths to be. He intuits that society's values are

not correct and then, rather than claiming that his *are* correct as some would-be

moral reformers have done, he adopts the "post-modern" position that no one's

are. Here moral skepticism is a meta-ethical reaction to an ethical problem, a

clash of values recast in metaphysical terms. Its motivations are fundamentally

political. This is the moral skeptic of popular lore, the *iconoclast*. This is the

moral skeptic whom you want to have at your dinner party, but whom you don't

want raising your children.

The other source of moral skepticism, the source of my own, is not so

colorful. It arises not out of hostility or indifference toward the values of others,

---

[1] Note that some people who are moral skeptics according to this definition would

be reluctant to call themselves "moral skeptics," e.g. Blackburn (1984, 1993a).

but out of an appreciation of the following difficult metaphysical problem. Most of us believe that certain moral claims such as (or similar to) "Lying is wrong," are true. The following question arises: *What makes such claims true?* Are they made true by the nature of the physical world? By facts about human psychology? By facts about what we would desire or believe if we were smarter, more rational, more imaginative, or better informed? By the meanings of the words used to express such claims? By some combination of these things? Or is there nothing to say at all about what makes such claims true beyond the fact that they just are?

I've concluded that this metaphysical problem has no solution because claims like "Lying is wrong," are never true. This makes me a moral skeptic, and I will defend this position in this chapter. The point for now, however, concerns my motivation in defending it. Whether or not the arguments I offer are ultimately convincing, I insist that one can doubt the truth of moral realism in the absence of any anti-social tendency, political agenda, or adolescent lack of seriousness. As noted in the last chapter, many philosophers meet moral anti-realism with hostility and an unwillingness to take it seriously. I believe this is because they mistakenly assume that anti-realism is inevitably allied with motivations and values they find unacceptable. The hostility and facile dismissiveness generated by this assumption can only undermine a clear-headed discussion of these issues, and I therefore request that readers to whom such attitudes come naturally do their best to put them aside insofar as this is possible.

2.1.2 Metaphysical v. Epistemological Skepticism

There is a long tradition of skepticism and anti-skepticism in philosophy. The skeptics typically win the battle, forcing their opponents to admit that, strictly speaking, they don't know what they ordinarily claim to know. The anti-skeptics typically win the war, forcing their opponents to admit that even they cannot take their own skepticism seriously. In light of this history, one might wonder if a defense of moral skepticism is not a lost cause. I believe that it's not. Some kinds of skepticism are more reasonable than others, and we can distinguish among them. More specifically, I suggest that we draw a distinction between *metaphysical* and *epistemological* skepticism.

Here's a typical skeptical fantasy. Most of us believe that there is a large landmass at the south pole called "Antarctica," a very cold place inhabited by penguins. A certain sort of skeptic might tell you that there is, in fact, no such place as Antarctica—just a grand conspiracy among historians, cartographers, politicians and travel agents bent on ensuring the rest of humanity's ignorance of this surprising fact. This skeptic claims that you cannot rule out the possibility that Antarctica is fictitious and that as a result you can't *know* that it exists.

As an anti-skeptic you might respond as follows: "Fair enough. According to the strictest standards of knowledge I don't *know* that Antarctica exists, but so what? You've given me no reason to take seriously your suggestion that it doesn't, and neither of us could take this suggestion seriously if we wanted to."

The epistemological skeptic's nagging question, "But how do you *know*?" is easily answered. The anti-skeptic can either admit that, strictly speaking, she has no "knowledge," and then downplay the importance of "knowledge" in this sense, or she can simply adopt a more relaxed standard for what counts as "knowledge" and thus stand behind her ordinary epistemic claims. Either way, the epistemological skeptic's point is trivialized, or at least put in its proper place.

While the epistemological skeptic asks, "How do you *know* that your belief in Antarctica is true?" the *metaphysical* skeptic asks a different question: "How can it *be* that your belief in Antarctica is true?" Put another way, the metaphysical skeptic asks, "If your belief is true, in what does the truth of your belief consist?"[2] On the one hand, the epistemological skeptic demands an explanation of how you *know* your belief is true, a demand that, strictly interpreted, is not so reasonable. The metaphysical skeptic, on the other hand, merely wants to understand what *makes it the case* that your belief is true if in fact it is. Unlike the epistemological skeptic's demand, this one is very easy to meet, so easy to meet that one is liable to misunderstand what is being asked. What would make your belief in Antarctica true is the instantiation of some physical state of affairs which includes the existence of a large landmass at the

---

[2] Instances of metaphysical skepticism are relatively rare. One example is the skepticism concerning meaning advanced by Saul Kripke in *Wittgenstein on Rules and Private Language* (1982).

south pole.[3] Maybe you can know that this belief of yours is true and maybe you can't, but it's not at all difficult to say what makes it true if it's true. While, the epistemological skeptic's enterprise amounts to an unreasonable demand that you rule out certain bizarre possibilities which, if actual, would render your beliefs false, the metaphysical skeptic's demand arises out of a perfectly reasonable expectation that one who believes something should be able to explain what it is that makes that belief true.

The moral skepticism advanced in this chapter is a form of metaphysical skepticism. The question with which we are concerned is not "How can you know for sure that your moral beliefs are true?" but rather, "How could it *be* that

---

[3] By casting the metaphysical skeptic's question as "What makes your belief true?" or "What makes it the case that your belief is true?" I do not mean to suggest that the metaphysical skeptic is necessarily inquiring after "truthmakers" in the narrow sense employed by Armstrong (1998). I do not mean to presuppose that all true claims are made true in each case by the existence of some *physical* object or state of affairs. As it happens, in the particular case of metaphysical Antarctic skepticism, the answer to the metaphysical skeptic's question does (or may) amount to the presentation of a truthmaker in Armstrong's sense. Nevertheless, I want to leave open the possibility that moral claims can be true without having Armstrongian truthmakers. I do insist, however, that in the case of true moral claims there must be some informative story to tell about what makes them true, at least insofar as such claims are not self-evidently true.

anyone's moral beliefs are true? What could make a moral claim true? In what does a moral claim's truth consist? If two people have a moral disagreement, what could make one of them correct and the other incorrect?" It's true that the moral skepticism I defend calls into question a wide range of common sense beliefs, but it does so by a very different means from typical, epistemological skepticism.   Unlike the epistemological skeptic, I'm not in the business of concocting bizarre but in-principle-possible scenarios that would, if actual, render our ordinary moral beliefs false. My worry is about the very *being* of moral truths under *ordinary circumstances*.

Whether or not these doubts about moral realism are well-founded, they shouldn't be dismissed as "mere skepticism." If we have moral beliefs we should be able to give some sort of account of what makes them true. Whether or not this can be done is an open question.

2.1.3 Moral Intuitions and Moral Skepticism

In ordinary, first order moral discourse we do not call into question all moral claims at once.   Rather, we question particular moral claims against a background of shared moral beliefs and/or values.[4] Ordinary moral discourse is

---

[4] Moral discourse tends to highlight our points of disagreement and minimize the appearance of common ground, but there is always *some* common ground, and sometimes quite a bit.

straightforwardly practical, aimed at guiding action in the here-and-now. As a result, those engaged in ordinary moral discourse have no reason to question the truth of uncontroversial moral claims. The fact that a moral claim *seems true* to all concerned is, for the purposes of ordinary moral discourse, enough to establish that it *is* true. Outside of philosophy, all moral claims are innocent until intuited guilty.

Moral skepticism is, roughly, the idea that at all moral claims are "guilty," not in the sense that they reflect undesirable values, but rather in that they are simply not *true*. Some moral claims have nearly universal intuitive appeal, for example: "It's wrong to kill people for fun." If we assume at the outset that our first order moral intuitions are our only guides to moral truth, then moral skepticism is a non-starter. We know that there are moral claims that reflect the values of all but a handful of deranged individuals. If (near) universal intuitive appeal is enough to make a moral claim true then we already know that there are at least some true moral claims and that moral skepticism is false.

To take moral skepticism seriously we must put our first order moral intuitions aside, at least insofar as they involve a commitment to realism. To conclude that there are moral truths on the grounds that there are a number of moral claims that appear to be true is to beg the question. Some may say that a question-begging response is a perfectly adequate response to moral skepticism. This brings us back to issues raised in the two previous sections.

I noted above that moral skepticism of the kind that arises out of a difference of values is really first order discourse recast as second order

discourse. The "political" skeptic's primary complaint is with other people's values, with their first order intuitions. Since this skeptic's issue is a first order issue, the standards of first order discourse may still apply. The skeptic says "Why?" and we reply "Why not?" If the skeptic's position is nothing more than an expression of his values, of his intuitions, then there is nothing wrong with simply reaffirming one's own values and intuitions in response.

But the skepticism we are considering is not first order moralizing in disguise. Our doubts apply equally to all moral claims, irrespective of their intuitive appeal or lack thereof. Still, one might wonder why intuitions are not enough. They're not enough because (a) our concern is with metaphysical moral skepticism, (b) because we recognize implicitly that metaphysical moral skepticism raises a legitimate question, and (c) because this question cannot be answered by appeal to moral intuition.

Recall that the metaphysical skeptic asks not "How do you know that your belief is true?" but rather "What makes your belief true?" Intuitions can provide an adequate response to the first question, but not the second. Consider the following analogy. Suppose you believe that Professor Moriarty is guilty as charged. The epistemological skeptic asks, "How do you know this?" You might respond "Because Sherlock Holmes said so." This response may be perfectly adequate, depending on the context. After all, Mr. Holmes is a celebrated sleuth, and his having concluded that such and such is the case is, under ordinary circumstances, evidence enough that it is. If the epistemological skeptic persists—"But how do you know Holmes isn't lying? How do you know that man

is really Holmes? How do you know he's not a computer generated phantom?"—his skepticism descends into triviality.

But now consider the metaphysical skeptic's question: "What makes it true that Moriarty is guilty?" To respond by saying "Because Sherlock Holmes said so," is no good. The fact that Sherlock Holmes said he's guilty may serve as good *prima facie* evidence for the claim that he's guilty, but it doesn't *make it true* that he's guilty. What makes it true that he's guilty is that he *did the deed*, and no one's opinion about the matter, Sherlock Holmes' included, is a substitute for the cold, hard facts concerning what took place at the scene of the crime.

Perhaps you think of your moral sense as something like your own personal Sherlock Holmes, a remarkable device that allows you to sniff out moral truths.   Whether or not you possess such a device, you must acknowledge that your having the moral intuitions that you have does not *make your moral beliefs true* any more than Holmes' concluding that Moriarty is guilty *makes him guilty*. This is what I mean by our implicit recognition that the metaphysical skeptic raises a legitimate question and that this question cannot be answered by appeal to moral intuition. We rely on our moral intuitions to guide our deliberations, but we are not so egocentric as to believe that our intuitions about what's right and wrong *determine* what's right and wrong.[5] I might oppose capital punishment, for example, because I feel that it's wrong, but my feeling that it's wrong is certainly not what *makes it wrong* if it is.

---

[5] Perhaps our moral intuitions *idealized* in some way could determine what makes things right or wrong. This is a possibility we will consider later.

Most people do not worry about what makes their moral beliefs true, but they do believe (if only implicitly) that *something* must make them true. Once again, people implicitly recognize that the metaphysical moral skeptic raises a legitimate question, even if they have no mind to try to answer it. There must be some story to tell about what makes moral claims true, and to tell this story we must put our moral intuitions aside since our intuitions are, at best, guides to moral truth and not moral "truthmakers."

2.2 Moral Realism and the Need for Fundamental Moral Principles

2.2.1 The Structure of Moral Principles

Typical moral claims predicate some moral property[6] of some set of persons, actions, or institutions. Some examples:

---

[6] Here I have in mind "abundant" rather than "sparse" properties, which is to say that properties in this sense are not restricted to those that may serve as part of a minimal basis on which all else supervenes.

Another issue concerning properties is whether or not they are more than abstractions from predicates. I intend to remain neutral on this issue. In what follows I will speak, for example, of the property of moral wrongness and argue that nothing has this property. However, I have no problem with using the term "property" slightly differently and saying that there simply *is* no property of moral

1. Slavery is immoral.

2. Lying is wrong.

3. Fred is immoral.

4. Abortion is morally permissible.

5. Killing your paternal uncle with a lead pipe in the conservatory on a Tuesday afternoon in order to more quickly receive the money he has left you in his will is wrong.

(Note that here and elsewhere "wrong" is short for "morally wrong.") Not all moral claims are of this form (e.g. "Lying is the moral equivalent of stealing"), but many are, and if no such claims are true, moral realism is all but empty.

We can understand each of these claims as stating a logical relation between two properties, and we can rewrite each of them in a way that makes this plain.

1. If X has the property of being an instance of slavery, then X has the property of being immoral.

2. If X has the property of being an act of lying, then X has the property of being wrong.

wrongness. Perhaps there are good reasons to favor one way of speaking over the other, but at present I have no settled opinion on the matter, only what amounts to a stylistic preference.

3. If X has the property of being Fred, then X has the property of being immoral.

4. If X has the property of being an act of abortion then X has the property of being morally permissible.

5. If X has the property of being an act of killing one's paternal uncle with a lead pipe in the conservatory on a Tuesday afternoon in order to more quickly receive the money he has left one in his will, then X has the property of being wrong.

The properties named in the antecedents of each of these claims are examples of what G. E. Moore (1959) calls "natural properties," which are opposed to "non-natural properties." Moore believes that most properties are natural properties, the everyday sorts of properties that are perceivable through the senses and/or investigated by science, but that moral properties, chief among them the property of being "good," are non-natural properties, perceivable only through a kind of intellectual moral sense. In Moorean terms, each of the claims above describes a logical relation between a natural property and a non-natural, moral property.

Non-natural properties are very much out of fashion these days, a welcome development to my mind. Nevertheless, it does seem that Moore was on to something with his distinction between natural properties and non-natural properties (where the latter include what we would recognize as moral properties). Some property ascriptions automatically commit one to value judgments. If, for example, we say that Katie's action has the property of being

wrong, then we've committed ourselves to an evaluative claim concerning what she did.[7] At the same time, it seems that we can ascribe some properties to persons, actions, etc. without making any kind of value judgment at all. If we say that Katie's action has the property of being an act of knowingly making an untrue statement with the intention of creating a false belief in another, we haven't yet passed judgment on what Katie did. Two people who disagree entirely about the merits of Katie's behavior can agree that her action has this property. Even if we say that Katie *lied*, we haven't, strictly speaking, passed judgment on Katie or her action. Whether or not it was at all wrong, bad, unjust etc. for her to lie in that situation remains, as Moore would say, an "open question."

Thus, it seems that some properties are clearly *evaluative* (the property of wrongness, for example) and that some properties are *value-neutral*, for example, the property of being an act of knowingly making an untrue statement with the intention of creating a false belief in another. That is not to say that the latter property is morally irrelevant. Indeed, if we hear that someone's action has this property our suspicions are rightly raised. The point is rather that no evaluative conclusion follows logically[8] from something's having this property,

---

[7] This follows trivially from the fact that to call something "wrong" in the relevant sense is to make a moral evaluation.

[8] By "logically" I have in mind what is sometimes called "broad logical implication," which allows for the use of additional analytic premises. For example, the fact that Joe is unmarried follows by broad logical implication from

whereas evaluative conclusions do follow logically from the fact that some action or person is wrong, bad,

For our purposes, the distinction between evaluative and value-neutral properties can serve as a surrogate for Moore's distinction between natural and non-natural properties, without the metaphysical baggage that comes with committing ourselves to the existence of non-natural properties. Thus, instead of thinking of the moral claims above as stating logical relations between natural and non-natural, moral properties, we can think of them as stating relations between value-neutral and evaluative properties. Because the property of being a lie is value-neutral (in our sense), two people who disagree entirely about the merits of Katie's behavior can still agree that she lied. Because the property of being Fred is a value-neutral property, two people who disagree entirely about *that guy's* (i.e. Fred's) moral worth can still agree on who counts as Fred. Likewise, the defenders of slavery were able to agree with their abolitionist opponents about what counted as slavery. To them—and to us, insofar as the term's meaning has not changed—"slavery" is a value-neutral term,[9] which is to

---

the fact that Joe is a bachelor, even though this inference makes use of the premise "all bachelors are unmarried."

[9] One might think that the term "slavery" cannot be value-neutral for a different reason than the one suggested here. Slavery might be defined as ownership of one person by another, and one might claim that ownership is a normative concept because to own something is to have certain rights with respect to it.

say that their evaluative disagreement was not a disagreement about what was or was not an instance of slavery. (It may seem strange to say that "lying" and "slavery" pick out value-neutral properties. Bear in mind, once again, that we are not saying that the presence or absence of these properties is of no evaluative import. Rather, the point is that it doesn't follow *logically* from the fact that something is a lie or an instance of slavery that it is wrong. One can claim that slavery is not wrong without *contradicting* oneself. Mounting a defense of slavery is not a *logical* mistake.) In contrast to "lying," "slavery," and the rest, the terms "wrong" and "immoral" are evaluative. Two people cannot disagree about which things have the property of being "wrong" or "immoral" without having some sort of evaluative disagreement.

The line between evaluative and value-neutral properties may not always be clear. Let us reconsider the property of being a lie and the property of being an instance of slavery. Some might say, contrary to what I asserted above, that you can't call something a "lie" or an instance of "slavery" without passing judgment on it. Again, while I agree that these terms have strong, even overwhelming, moral overtones, I maintain that there is no evaluative component to these words' *meanings*, that it is not a *logical contradiction* to say that it was right to tell a certain lie or that there could be an instance of slavery that is not wrong. But others, especially those who subscribe to the dictum "meaning is use," may disagree. Some further examples: One might argue that by calling an

---

This particular problem can be avoided by defining slavery in terms of possession rather than ownership.

act "courageous" we have committed ourselves to some sort of evaluative conclusion, that we would contradict ourselves by saying of some act that it was both courageous and worthy of no praise whatsoever. When we say that someone has a "right" to free speech, we might be making a value-neutral claim about the content of the relevant civil laws, but, alternatively, we might be making an evaluative claim, a claim about how things ought to be. For our purposes, we can live with these ambiguities by adopting a conservative policy. We'll say that the value-neutral properties are, for our purposes, those that are uncontroversially value-neutral in the relevant context.

Once again, the moral claims listed above ascribe evaluative properties to persons, actions, and institutions. Each of the numbered claims above can, if necessary, be recast as a statement connecting a value-neutral property to an evaluative property. If "lying" and "slavery" are not, in the sense we have defined, value-neutral terms, then we can substitute "making untrue statements with the intention of creating false beliefs in others" for "lying" and "the possession of one person by another" for "slavery," thus making the claims above fit the paradigm.

Let's call claims of the form "If X has such and such value-neutral property, then it has such and such moral property," *moral principles*. As noted above, there are claims which would rightly be called "moral principles" that are not of this form, but this is a harmless simplification for our purposes. We're raising doubts about moral realism, and any kind of moral realism worthy of its name requires that there be at least some true moral principles in the sense we

just defined. At least some claims like "Abortion is wrong" or "Abortion is morally permissible" or "Killing people for fun is wrong" must be true if moral realism is going to be true. Because moral realism requires the existence of true moral principles, we can raise doubts about moral realism by raising doubts about moral principles.

Our skeptical question, slightly refined, is this: What could make moral principles true? Put another way: What establishes the connections between the value-neutral properties that define the objects of moral interest (e.g. the property of being an act of abortion) and the evaluative properties that, according to moral realism, inhere in those objects (e.g. the property of being wrong). In what follows we will consider the relationship between value-neutral and evaluative properties and the various means by which they might be connected.

2.2.2 The Supervenience of Value and the Necessity of Moral Principles

Suppose, for example, that I tell you that Giulio is a skilled singer. You might then ask, "What makes him a skilled singer?" Now suppose I say, "Nothing. He's just a skilled singer." You'd be puzzled, no doubt. "But surely," you might say, "there must be something about the way he sings that makes him one among the skilled singers. Is it his ability to project his voice? His ability to accurately hit a wide range of notes? What?" And then I respond, "Oh, no, no. Nothing like that. He's just a skilled singer, that's all." This is nonsense. Giulio

can't be a skilled singer without there being some further story to tell about the features of him and his singing that make them skilled. Giulio has the property of being a skilled singer in virtue of having some combination of more basic properties, properties concerning his ability to project his voice, to accurately hit a wide range of notes, etc.

The same goes for moral properties. If some action is wrong, it can't just *be* wrong. It has to be wrong for some reason, because that action has certain properties that *make it wrong*. Perhaps it's an action that fails to maximize utility, or an action based on a maxim that cannot be universalized, or perhaps something else entirely, but one way or another there must be some story to tell about the properties of that action in virtue of which it has the further property of being wrong.

A more precise way to capture this idea is in terms of *supervenience*. Given two classes of properties (A and B), we say that the B-properties supervene on the A-properties if and only if it is impossible for two objects, not necessarily in the same world, to differ in their B-properties without also differing in their A-properties. (The B-properties are the "supervening properties" and the A-properties are the "subvening properties.") A computer screen makes for a nice illustration of the supervenience relationship. We can describe (say true things about) the state of a computer screen at different levels. At a high level of description, we can say of a given computer screen that it depicts a dog, or a sunny day,   Alternatively, we can describe it at a lower level by talking about its pixels. We can say that the screen is red at point 234-317 and green at point…

The image-level properties supervene on the pixel properties. How so? Recall our definition of supervenience, which requires that no difference in B-properties is possible without a difference in A-properties. If the screen depicts a dog, it won't cease to depict a dog without some of its pixels' changing. And for the same reason, if two computer screens are identical pixel for pixel and one of them depicts a dog, then the other must depict a dog as well.[10] An easy way to think of the supervenience relation is in terms of "fixing." The pixel properties of a computer screen "fix" its image properties. Setting the pixels sets the image, but not the other way around.

Just as the higher-level, image properties of a computer screen are fixed by its lower-level pixel properties, the higher-level moral properties of actions, persons, institutions etc. are fixed by their lower-level properties. In other words, moral properties supervene on lower-level properties. But note that these lower-level properties, if we go low enough, are *value-neutral*. This becomes clear

---

[10] Note that the supervenience relationship is (almost always) asymmetric. (Why "almost?" A consequence of the definition of supervenience used here is that a set of properties will (trivially) supervene on itself and therefore exhibit a symmetric supervenience relationship.) You could change a few pixels on a screen showing a dog and it might, probably will, continue to show a dog. Likewise, you could have two screens that are identical as far as what they are depicting at the level of images (i.e. at the level of dogs and sunny days) but different in their pixels. For example, one screen might have a higher resolution than the other.

when we consider the fact that moral properties supervene on low-level, physical properties. Two worlds that are identical in their low-level physical properties, that are matched atom for atom, are going to be morally identical as well.[11] Likewise, two actions, persons, etc. that differ in their moral properties are going to have to be different in some low-level, physical way. Such differences may be subtle and may be relational, but they have to be there.

Thus, moral properties, and evaluative properties more generally, supervene on value-neutral properties. Before we consider the implications of this conclusion, we should consider an assumption we've made along the way. We've assumed that low-level physical properties are value-neutral. *Prima facie*, this assumption may appear unassailable. When we say that some object consists of such and such molecules arranged in such and such way, we hardly seem to have committed ourselves to any sort of value judgment about that object or anything else. Nevertheless, while low-level property ascriptions taken in isolation or in small clusters are undeniably value-neutral, perhaps, if we get enough of them together, they will become evaluative. How? Bear in mind our stipulation that a property is evaluative if by ascribing it to something we logically commit ourselves to some evaluative claim. Imagine that we have a complete low-level, physical description of our universe. It might be possible to logically

---

[11] Should it turn out that some morally relevant mental properties fail to supervene on physical properties, we can restrict our notion of supervenience to nomological supervenience. In suggesting this I am assuming that mental properties supervene nomologically on physical properties.

deduce from this description all sorts of high level facts about our world, and maybe—we don't want to rule it out at this point—we could even reach some evaluative conclusions by such means. That said, it's still the case that moral properties supervene on low-level physical properties, even if we can't be absolutely sure that some exceedingly complex low-level physical properties are not in fact evaluative properties according to our definition. For now we'll simply ignore this potential wrinkle and say that all low-level physical properties are value-neutral and that therefore all moral properties supervene on value-neutral properties. (We'll return to this issue in Section 2.4.5.)

What does this tell us? In the previous section we concluded that moral realism depends on the existence of moral principles, claims that connect value-neutral properties to moral properties. The supervenience of the moral on the value-neutral explains why this is so. What moral properties a thing has is determined by[12] the more basic, value-neutral properties that thing has, and, at the most basic level, by its low-level physical properties. This supervenience relationship has another important implication. Some of these moral principles are going to have to be true *necessarily*.[13]

This necessity falls out of the nature of moral supervenience. Once you've set the pixels on a computer screen, the on-screen images *must* be

[12] At the very least this determination is nomological. See previous footnote.

[13] For all we've said so far it's possible that some or all of these principles will be infinitely complex. Nevertheless, such principles, however unfathomable, will be true necessarily.

whatever they are going to be. Likewise, the fact that moral properties supervene on lower-level, value-neutral properties means that once those lower-level properties are in place, the moral properties *must* be whatever they will be. For example, if it's true that lying is wrong, then if an action has the property of being a lie, that action must have the property of being wrong. Of course, "Lying is wrong" may not be such a plausible principle. Presumably there are some instances in which it's not wrong to lie. But this point about the necessity of moral principles remains even if overly simple principles such as "Lying is wrong" are never true. Assuming moral realism is true, there has to be some fact of the matter about when lying is or is not wrong, and that means that there must be some more complicated principle of the form "Lying is wrong except in such and such cases," where the "such and such cases" are spelled out in terms of lower-level, value-neutral properties at whatever level of detail is necessary to get things right.[14] Regardless of what those details turn out to be, this principle will be true necessarily. If something has the value-neutral property of being a lie and does not have the value-neutral properties that would make it an exception to the rule "Lying is wrong" then it *must* have the property of being wrong.

---

[14] Moral realism might allow for certain moral indeterminacies. Perhaps when it comes to lying, for example, there will be cases in which it is indeterminate whether or not lying is wrong. The point for now, however, is that moral realism requires that *some* moral claims be determinately true. If nothing is determinately wrong—not even, say, torturing people for fun—then it would be a bit much to call the correct meta-ethical view "moral realism."

We defined moral principles as those claims that link value-neutral properties to moral properties. If we keep this definition, not all moral principles will be necessarily true. "Fred is immoral," is, according to our definition, a moral principle because it links the value-neutral property of being Fred to a moral property. But unless it is essential to Fred's being that he is immoral—and I think we should give him the benefit of the doubt—then this claim about him is not true necessarily, even if it's true.[15] But "Fred is immoral" isn't the sort of thing we have in mind when we think of moral principles anyway. The relevant moral principle seems to be "If one does such and such things, then one is immoral," which, combined with the contingent fact that Fred did those things, yields the claim above. And this principle, whatever it turns out to be, is necessarily true. (One can't have the kinds of value-neutral properties that make one immoral without being immoral.) Of the claims that link value-neutral to evaluative properties, it seems that it's the necessarily true ones that do all the work. The other ones like "Fred is immoral" are really just consequences of the necessarily true ones and some contingent facts. Furthermore, claims like "Fred is immoral" couldn't be true unless there were some necessarily true moral principles linking the lower-level, value-neutral properties, in this case the ones that Fred happens to have, to the property of being immoral (or virtuous, just, unjust, etc.).

---

[15] If it follows logically from the fact that something is Fred that that thing is immoral, then "Fred" is not a value-neutral term and as a result "Fred is immoral" is not a moral principle. The thought here, however, is that Fred's immorality might be an *a posteriori* necessity.

Therefore, from now on when we speak of moral principles, we are speaking of those necessarily true claims that link low-level, value-neutral properties to moral properties.

Moral realism requires such principles. Without them there can be no story to tell about the why the various things that are right, wrong, immoral, just, unjust, etc. have those properties.

2.2.3 Fundamental Moral Principles

So far we've determined that if moral realism is true there must be some set of true moral principles, where these moral principles are necessarily true claims that connect the moral properties to the lower-level, value-neutral properties on which they supervene. Call this set of principles the *true moral theory*. Our skeptical question, recast once again: What makes this theory true?

A question of the form "What makes X true?" may be understood and answered in a number of different ways. One type of answer we've covered already. We might read such questions as asking for a lower-level description of some contingently existing state of affairs. For example, when one asks what makes Giulio a skilled singer we can launch into a more detailed discussion of his abilities. But this can't be the sort of answer we're looking for here. "Giulio is a skilled singer" states a contingent fact, one that is analogous to "Fred is

immoral,"[16] but the moral principles with which we are concerned are supposed to be necessarily true. A contingent fact can't explain what makes a necessarily true claim true since the necessarily true claim would have been true even if the contingent fact failed to obtain.

One might answer a "What makes X true?" question by giving a causal explanation: "What makes the Sun shine?" (More awkwardly: "What makes it true that the Sun shines?") Answer: "Nuclear Fusion of hydrogen into helium in the Sun's core." But this can't be the sort of answer we're looking for either since causal explanations are just a species of explanations that appeal to contingent facts. Besides, it should be clear that if lying is wrong, lying didn't get to be wrong as the result of some physical process.

If we aren't going to explain what makes moral principles true by appeal to contingent truths, that leaves only necessary truths. Let's begin by dividing the necessary truths into moral truths and non-moral truths. We'll say that moral truths are any truths that make claims about evaluative properties. (There may be evaluative properties that are non-moral, e.g. aesthetic properties, but for our purposes it won't hurt to throw them in with the moral properties.) Moral principles, as we've defined them, are a subset of the necessary moral truths. Necessary moral truths that are not moral principles in our sense include moral tautologies (e.g. "If X is wrong then X is not right") and other necessarily true moral claims that are not of the form, "If X has such and such value-neutral

---

[16] Here I am assuming once again that Fred is not necessarily immoral.

properties, then X has such and such moral properties" (e.g. "Lying is the moral equivalent of stealing.")

Non-moral necessary truths are more of a mixed bag. There are truths such as "All bachelors are men," typically thought of as analytic and true by dint of meaning. There are mathematical truths, typically thought of as synthetic and *a priori*, and there are truths such as "Water is H_2O," thought by many to be synthetic and (necessary) *a posteriori*. Other candidates include laws of nature, principles of mereology, principles of modal metaphysics, and whatever non-moral necessary truths can't be known at all. Perhaps there are other categories that belong on this list, but these are the standard ones.

Could an appeal to non-moral necessary truths explain what makes a moral principle true? Could we explain what makes "Lying is wrong" (or some other moral principle) true by appeal to the fact that "All bachelors are men," or that "Two plus two equals four," or that "Water is H_2O" or by appeal to some other *non-moral* necessary truth? I highly doubt that non-moral facts of any kind, necessary or contingent, are going to explain what makes moral principles true. I have no argument for this. I can only leave it as a challenge to those who disagree to explain how this might be. For now I'll assume that this is right, that we're not going to understand what makes moral principles true by appeal to non-moral necessary truths, i.e. by appeal to necessary truths that, when expressed in English, make no reference to moral properties.[17]

---

[17] This simple formulation suggests the right idea, but is not strictly adequate. Its inadequacy is a special case of a broader problem concerning the "autonomy of

ethics" (Prior, 1960; Kurtzman, 1970; and Jackson, 1974) which may be illustrated as follows: Let p be the claim that two plus two equals five. Let q be the claim that two plus two equals four. Let m be the moral principle that lying is wrong. Suppose that $\sim p$, q, and m are all necessarily true. Let d_1 be the disjunction of m and p. Let d_2 be the disjunction of m and q. It seems clear that d_1 is a moral principle if and only if d_2 is a moral principle. Suppose that neither is a moral principle. From d_1 and $\sim p$ (which is also not a moral principle), we can derive m. Suppose instead that d_1 and d_2 are moral principles. We can derive d_2 be from q, which is not a moral principle. Either way we have derived a necessarily true moral principle from one or more non-moral necessary truths.

I suggest that we reject the assumption that d_1 is a moral principle if and only if d_2 is. While these claims have a common logical structure (moral claim disjoined with non-moral claim), they are very different in terms of how much they tell us about the nature of moral reality. By disjoining a moral claim with a claim that is necessarily true, we say nothing about the nature of moral reality. It's rather like pairing you up with an omniscient Trivial Pursuit partner. Your team's answers tell us nothing about your knowledge because your partner can carry you all the way if necessary. In contrast, by disjoining a moral claim with a claim that is necessarily false we make a bold moral statement, much as your team's answers do reflect your knowledge when you've been paired with a complete ignoramus.

Thus, in response to this particular objection we can say that moral claims disjoined with necessarily false non-moral claims count as moral claims and that

That leaves only moral necessary truths. Could a moral necessary truth explain what makes a moral principle true? Absolutely. The fact that "Lying is wrong" might very well explain the fact that "Lying on Tuesdays is wrong." More substantively, a Kantian might say that the fact that "Treating others as a means only (which is necessarily what happens when one lies) is wrong." If there are true moral principles, it should come as no surprise that some of them are explicable in terms of others. Bear in mind, however, that we moral skeptics are calling into question *all putative moral principles at once*. When presented with the aforementioned explanation for why it's true that lying is wrong we are not going to be satisfied until we hear an explanation for what makes the above Kantian principle true. We can explain what makes some moral principles true by appeal to other moral principles, but we can't explain what makes all the moral principles true by appeal to *any* moral principles. Such an explanation would be circular.

What about moral necessary truths that are not moral principles as we've defined them, principles such as "Lying is morally equivalent to stealing," and "If X is wrong, then X is not right?" I fail to see how one could explain what makes "Lying is wrong" or any other moral principle true by appeal to claims such a these without also appealing to moral principles in the sense we've defined (e.g. "Stealing is wrong"). Remember that a moral principle is a claim that forges a

---

moral claims disjoined with necessarily true non-moral claims do not. There are, no doubt, variations on this objection that are not handled by this response, but I'm content to assume that variations on this response can meet them.

link between a moral property and the value-neutral features in virtue of which things have that property. It's hard to see how moral claims like those above which do nothing to bring moral properties and value-neutral properties together can explain what makes the principles that do accomplish this true. Here, too, I can only issue a challenge to those who disagree.

Consider where this leaves us in terms of our goal of explaining what makes the true moral theory true. We can't appeal to contingent truths at all because the truths we want to explain are necessary, nor can we appeal at all to moral principles since to do so would be circular. That leaves us with non-moral necessary truths and necessary moral truths that are not, in our sense, moral principles. Neither of these look very promising, and they don't look any more promising in combination. Some moral principles can be explained by appeal to other moral principles, but that still leaves those other moral principles unexplained. We are forced to conclude that if moral realism is true, some moral principles must be true in spite of the fact that there is *no explanation for what makes them true*. They must be *brutely* true. Call these moral principles the *fundamental moral principles*.

That moral realism requires the existence of fundamental moral principles may come as no shock to its defenders. It's worth emphasizing, however, that what we mean by "moral principles" is somewhat more narrow than what one might otherwise have in mind. The fundamental moral principles have to have a very specific form. They must reveal the nature of the supervenience of the moral on the value-neutral by telling us what value-neutral properties give rise to

what moral properties, and, if they are not to be mysterious, they must do so in such a way that we are not left asking "Why is that so?" Because the fundamental moral principles are fundamental, they can admit of no further explanation.

2.3 The Nature of Fundamental Moral Principles

2.3.1 The Variety of Moral Principles: General and Specific, Direct and Indirect

Moral realism depends on the existence of fundamental moral principles, principles that determine which value-neutral properties give rise to which moral properties and that cannot be explained in terms of other facts. In considering whether or not there could be such principles, we will divide all possible moral principles into sub-groups and handle these sub-groups separately.

First, moral principles can be more or less *general*. Take our favorite moral principle: "Lying is wrong." This principle is fairly general, but also fairly specific. It allows one to draw moral conclusions about a great many actions, namely all acts of lying. At the same time, though, there are many actions about which it says nothing at all. For example, it says nothing about the moral properties connected to acts of stealing.

The most general principles are ones that, provided that we have enough contingent information, allow one to draw moral conclusions about all actions. A

standard utilitarian principle is a good example: "An action is wrong if and only if it fails to maximize the overall sum of happiness." Near the other extreme are very specific principles, for example: "Killing your paternal uncle with a lead pipe in the conservatory on a Tuesday afternoon in order to more quickly receive the money he has left you in his will is wrong." But even this principle is not maximally specific, as it allows one to draw moral conclusions about a number of possible actions.

The shape of the moral landscape, i.e. the content of the true moral theory, will be determined in part by the level of generality that the most general true moral principles have. A moral theory will be rather "theoretical" if it contains some very general moral principles, principles from which one can derive a wide range of moral conclusions (given the relevant contingent information). If, for example, the aforementioned utilitarian principle is correct, then the correct moral theory is maximally "theoretical,"; all of the moral truths follow from one fundamental moral principle. If utilitarianism (or some other maximally theoretical theory) is correct we can compare the moral landscape to a mosaic the tiles of which are all placed according to a single simple formula.

More moderately "theoretical" moral theories are possible. Instead of there being a single fundamental moral principle there might be several such principles, a set of "moral axioms."[18] Together these axioms would form

---

[18] Of course, any set of axioms can be conjoined to form a single axiom. The assumption made above is that each axiom will be as logically simple as possible.

something like a moral legal system, a set of consistent rules from which all the facts about how one must behave follow (once again, when supplemented with the relevant contingent information). If such a moral theory were correct then the moral landscape would be analogous to a mosaic with tiles placed according to some relatively small set of rules.

At the other extreme are moral theories that are not at all "theoretical." Each such theory is composed of a set of highly specific moral principles such as the one above about killing one's uncle. Such views are forms of what one might call *particularism* since, on this view, all the moral truths are particular. They are also all fundamental in the sense that they are not derivable from more general moral principles. We might call the view according to which all of the moral truths are maximally specific *extreme particularism*. If extreme particularism is true then for every action there will be a set of fundamental moral principles that specifically determine its moral properties, i.e. connect its value-neutral properties to its moral properties. According to this view the moral landscape is analogous to a mosaic with tiles placed according to no general rules at all. To give such a mosaic a complete characterization one must describe it tile by tile.[19]

---

[19] Some philosophers may consider themselves realists but deny that there is a true "moral theory." However, insofar as such philosophers accept the supervenience of moral properties on some set of lower-level properties, they are still committed to the existence of a true moral theory in our sense. In our terms, what they reject is a true moral theory that has a "theoretical" character, one that has true *general* moral principles.

Note, however, that complete characterizations are not the only kinds of characterizations of interest. One could have a mosaic the nature of which is determined by no rigid set of rules, but which nonetheless exhibits a number of interesting general patterns. It might be the case, for example, that most of the tiles in the upper-right quadrant are black, or that there are almost never more than four red tiles in a row, or… One might be able to make any number of useful generalizations about such a mosaic, even if there are no exceptionless principles of any generality that describe its design to the last tile. I mention this possibility because, to many, and especially to philosophers who have tried their hand at principled casuistry, this seems to be the right picture of the moral landscape.[20]

Moral principles, in addition to varying in their level of generality (contra specificity), also vary in the extent to which they wear their content on their sleeves. On the one hand, *direct* moral principles allow one to move relatively easily to moral conclusions about particular actions, persons, By knowing them one gets a relatively clear idea of what particular sorts of things are right, wrong, "Lying is wrong" is a direct moral principle. All one needs to know in order to apply this principle to a particular case is whether or not the action in question is a lie. The utilitarian principle mentioned above is somewhat less direct, but still rather direct. We can never be sure of how some action will affect the overall sum of happiness, but we can often make a pretty good guess, at least in the short term.

---

[20] I believe that this was G. E. Moore's (1959) view.

On the other hand, *indirect* moral principles are ones from which it is very difficult, if not impossible, to draw conclusions about particulars, at least when it comes to controversial matters. Consider the following indirect principle: "An action is wrong if and only if we would disapprove of it if we were fully rational and fully informed."[21] This principle, like all indirect principles, tells us almost nothing about the actual content of the true moral theory, which tiles fall where, so to speak.

2.3.2 Fundamental Moral Principles: General or Specific?

Could there be fundamental moral principles that are specific? Here's one that should be as good as any other: "Lying in order to get someone to give you her money so you can spend it on luxuries for yourself is wrong." This principle is very plausible in the sense that it fits in well with our moral intuitions. I can't imagine that too many moral realists would want to say that this principle (or some variation on it) is not true. But our question is not whether it is plausible *per se*, but whether it is plausible *as a fundamental moral principle*. Could such a principle be true without there being some further explanation for why it's true? Wouldn't there have to be some story to tell about what the essential features of this action are that make it wrong, about the features it shares with other wrong

---

[21]This is a modified version of a principle put forth by Smith (1994). We will encounter it again in our discussion of analytic naturalism.

actions? Wouldn't there have to be some more general principle at work that explains why this sort of thing is wrong?

I have a hard time imagining how such a principle could be *brutely* true, how there could be no further explanation for why it's true. Moreover, it seems that in the absence of some more general moral principles, the only way that we could come to know such specific moral principles is through a Moorean extra-sensory moral intuition.

The above principle is specific and direct. Strictly speaking, there could be specific and indirect moral principles, but these would be very strange. An example would be "Lying in order to get someone to give you her money so you can spend it on luxuries for yourself is wrong if and only if we would disapprove of it if we were fully informed and fully rational." Principles such as these strike me as no more likely—and, if anything, less likely—to be fundamental moral principles than direct, specific moral principles.

In light of the above, I take it that specific moral principles, whether direct or indirect, can't be *fundamental* moral principles.[22] This means that the fundamental moral principles will have to be general, though they may be direct or indirect. Furthermore, the fundamental moral principles will either be *analytic*

---

[22] How specific is too specific? The fundamental moral principles have to be at least less specific than the specific principles described above. This does leave room for the possibility of fundamental moral principles that, though somewhat general, are not extremely general. "Lying is wrong," for example, is still a viable candidate.

or *synthetic*. That is, they will either be true by dint of meaning (e.g. "All bachelors are men") or they will not be so. The naturalistic view according to which the fundamental moral principles are analytic is *analytic naturalism*. The naturalistic view according to which the fundamental moral principles are synthetic is *synthetic naturalism*. We will consider these two views in the sections that follow.

2.4 Analytic Naturalism

We've concluded that moral realism depends on the existence of fundamental moral principles that are general and brutely true. Very few claims are plausibly taken to be brutely true, true without any further story to tell about what makes them true. Some analytic claims do have this feature.[23] When presented with

---

[23] The most familiar analytic claims such as "All bachelors are men" have this feature, but not all of them. For example, some analytic claims are theorems of mathematical logic the proofs of which are long, difficult, and informative. Such proofs certainly count as "further stories to tell" about what makes those theorems true. It strikes me as extremely unlikely, however, that there will be analytic fundamental *moral* principles the truth of which can be explicated by long proofs.

Might the fundamental moral principles be unobvious analytic principles of some other kind? This is the analytic naturalist's hope, and I know of no way to

such a claim we can explain what makes the *sentence* that expresses it true by explaining the meanings of the words it contains, but once we've gotten past the language and have grasped the meaning of the claim itself—the proposition, the non-linguistic truth-bearer—there is nothing left to explain about why it's true. One who understands English has no need to ask what makes it true that all bachelors are men.

The thought behind analytic naturalism is that the fundamental moral principle(s) is/are analytic. If that's so, then our skeptical problem is solved; we'll have fundamental moral principles that require no explanation. The question, then, is whether or not there are any plausible fundamental moral principles that are analytic.

2.4.1 Direct Analytic Moral Principles and Moore's Open Question Argument

Moore (1959) believes that his "open question argument" shows that moral properties could not be identical with natural properties, but what his argument actually shows is more modest. It shows, or at least very strongly suggests, that no *direct moral principle* can be *analytic*. Take the principle "lying is wrong." Given that X is an act of lying, it is still an "open question" whether or not X is

---

deprive her of it. The best we can do, I think, is to examine the proposals that are on the table and gauge our expectations based on what we find. That is the business of this section.

wrong. This doesn't mean that this principle can't be true. It just means that it can't be *analytic*. Were it built into the meaning of "lying" that all acts of lying are wrong, the question as to whether or not X is wrong, given that it's an act of lying, would not be open. Similar comments apply to the utilitarian analysis of "good." Given that X is an act that maximizes utility, it is still an open question whether or not X is good. Thus, "An act is good if and only if it maximizes utility" cannot be analytic. Moreover, it seems that similar comments will apply to any direct moral principle. Of course, this is an argument from example, and it's always possible that someone will find a counter-example, but in the hundred or so years since Moore's argument appeared, no one, to my knowledge, has produced a direct moral principle that appears to be true by dint of meaning. I think it's safe to conclude at this point that if analytic naturalism is true, it will have to come in the form of a theory with fundamental moral principles that are indirect.

Moore rejected all naturalistic analyses of moral terms, but he never explicitly considered such analyses in the form of *indirect* moral principles.

Since Moore's time a number of philosophers have formulated indirect moral principles that are sufficiently general to play the role of fundamental moral principles, that produce no apparent conflicts with first order moral intuitions (by virtue of their indirectness), and that are plausibly taken to be analytic. Such principles hold the promise of undercutting Moore's open question argument and placing naturalistic moral realism on a firm foundation. We turn now to a discussion of such principles.

Michael Smith (1994) offers an analysis of moral rightness that might serve as a fundamental moral principle. Here we'll consider a slightly modified version, one we've encountered already:

> An act is wrong if and only if we would disapprove of it if we were fully informed and fully rational.

This principle connects two properties: the moral property of wrongness and the property of being an act of which we would disapprove if we were fully informed and fully rational. The property of wrongness is a paradigmatic moral/evaluative property. If this moral principle is going to do the job of connecting moral properties to their subvening value-neutral properties, then the second property must be value-neutral. Is it? Philosophers typically think of "rational" as a normative term, which suggests that it is not. If it's not, then the above principle won't do. Nevertheless, we'll make the *conservative* assumption that this second property is value-neutral. Friends of this principle, and rationalism in general, may take offense at this suggestion, but none is intended. By making this assumption we're simply agreeing to give this principle a shot at solving our skeptical problem. If we can accept the above principle as true (and true by dint of meaning), we skeptics will pack up and go home without worrying about how

the property of being something of which we would disapprove if we were fully informed and fully rational supervenes on (other?) value-neutral properties.

The question is whether or not "wrong" could really mean "something of which we would disapprove if we were fully informed and fully rational." I don't see how it could. The concept of moral wrongness was alive and well long before we Westerners developed our present fondness for rationality. When ancient Israel's ruling elite declared that it's wrong for non-priests to enter the inner sanctuary of the holy Temple, they were not making a claim about idealized rational behavior. Of course, they may have had nothing against rationality. Perhaps they thought that more information and more rational appreciation of that information is more likely to push one away from sin than towards it, but that's not the same thing as their meaning by "wrong" "that of which we would disapprove if we were fully informed and fully rational."

Smith's analysis has some plausibility because people in many, perhaps most, cultures have attempted to reason about right and wrong. With this in mind, one might then jump to the conclusion that all people implicitly recognize that right and wrong is ultimately about nothing more than behaving rationally, but that conclusion is unwarranted. There is an important difference between being willing to reason about some domain of fact and believing that the facts in that domain are simply facts about what one would conclude if one reasoned properly. Take the belief that AIDS is caused by a virus. One might reason one's way to this conclusion. (In fact, some people did.) But that doesn't mean that the fact that that AIDS is caused by a virus is just the fact that we would

believe that AIDS is caused by a virus if we were fully informed and fully rational. Of course, it follows from the fact that AIDS is caused by a virus that *fully informed* people will believe that this is so, but to identify the fact that AIDS is caused by a virus with the fact that fully informed (and fully rational) people will believe this is to miss the point. What *makes it the case* that AIDS is caused by a virus is the obtaining of a certain state of affairs involving a virus and a disease. Therein lies the fact. That fully informed people will acknowledge this fact is just a trivial consequence of the fact itself. Note, however, that Smith does not take "fully informed" to mean fully *morally* informed, as this would trivialize his analysis. As a result, one cannot *assume* that fully informed and fully rational people could never approve of something that is wrong. This rather ambitious claim requires an argument, and my point at present is that an appeal to our willingness to reason about moral matters is not enough. One can reason about a domain of fact while believing that those facts are true in virtue of things that have nothing to do with us and our rationality.

But even if we were to grant that a commitment to reasoning about moral matters is enough to support Smith's analysis, it's unlikely that competence with moral concepts and terms requires such a commitment. In most times and places morality has fallen within the province of religion. Some religions have developed scholarly traditions according to which moral principles are discussed in a rational way, but not all of them have, and at the end of the day many, if not most, of these traditions take right and wrong to be grounded in some kind of higher metaphysical order, the comprehension of which may or may not be

accessible to human reason. Indeed, many traditions emphasize the importance of *faith over reason*. Want to avoid wrongdoing? Don't think too hard. Don't *intellectualize* your moral life. Just put your faith in God and He will guide you. The people who adopt this approach are clearly not thinking that what they're aiming for in avoiding wrongdoing is a life that is ideally rational.[24] Are such people confused about the meaning of "wrong?" The fact of the matter is that outside of Anglo-American philosophy departments "wrong" has a lot more to do with the Will of God than with idealized rational behavior. Bearing this in mind, it's hard to imagine that rationalism could be built into the meaning of "wrong" and other moral terms.

Another way of making the same point: Suppose a would-be philanderer asks, "Why shouldn't I?" One response: "Because it would be *wrong*!" Another response: "Because it's *something of which we would disapprove if we were fully informed and fully rational!*" There may be something to be said for the persuasive power of the second response, but the first one packs a lot more

---

[24] That God tells you not to do something may give you *a reason* to refrain, but this is not sufficient for moral rationalism. Everyone agrees that people have reasons to do things. The difference between rationalists and everyone else is that rationalists believe that Reason, i.e. *reason alone*, tells us what to do, that reason alone provides us with categorical rather than merely hypothetical imperatives. This rather esoteric idea was the brainchild of Immanuel Kant, and he knew it. His project was decidedly *not* one of giving expression to the meta-ethical thought of his contemporaries, the ancient Israelites, or anyone else.

punch, at least for most people. Why? Why should the first response be so much more forceful if these two responses *mean the same thing*? Perhaps the difference is merely one of rhetorical effectiveness—it's a punchier way of putting the point. That seems off the mark. Rather, there seems to be a genuine difference in content between these two assertions. The first response tells you that there is an *unconditional prohibition* of such behavior, that you are not supposed to do it regardless of what your interests or commitments are. The second response tells you that certain concerns—concerns which may or may not be terribly important to you—speak against such behavior. To say that something is wrong is to make a kind of "ultimate" condemnation.[25] Insofar as we are interested in observing the requirements of morality, our discussion is over as soon as we agree that the action in question is wrong. In contrast, to say that something would be disapproved of by us if we were fully informed and fully rational is to point out a potentially undesirable feature of the action in question.[26]

---

[25] Not "ultimate" in the sense that what is said to be wrong must be *gravely* wrong, but rather in the sense that it's wrongness does not depend on the interests or commitments of the agent in question.

[26] What I've presented here is essentially Moore's open question argument. It shows that Smith's principle must be either synthetic or unobviously analytic. The arguments presented above suggest that it is not analytic, obviously or unobviously. Once again, it's hard to imagine that all (or even most) competent users of the term "wrong" have an implicit commitment to the sort of rationalistic metaphysics that Smith has attempted to pack into the meanings of moral terms.

The concept of "rationality" is no doubt rather elastic, but its elasticity won't help in this case because Smith's rationalism requires a notion of rationality that is simultaneously very "thick" and very "thin" (impossibly thin, in fact). For something to be morally wrong it has to be unconditionally prohibited. It must follow from the fact that some act is wrong that it is "not to be done" regardless of the commitments or interests of particular people. This calls for a thin notion of rationality. If we build into the idea of rationality that rational people will behave in certain ways, behave as "reasonable" people do, then rationality, and ultimately morality, becomes optional. One can coherently ask, "Why must I be rational (and therefore moral) *in that sense*?" At the same time, Smith's rationalism requires a thick notion of rationality. If being rational amounts to little or nothing more than being logically consistent, then there are no substantive aims that "we" would share once we've been fully informed and made fully rational. There are among us selfish jerks, and selfish jerks can be fully informed and fully logically consistent. One might hope that there would *just happen to be* no selfish jerks among our fully rational and logically consistent counterparts, but even that would not suffice for reasons that will be explained in the next section. As it happens, this is a moot point because the question before us is whether or not there is a sense of "rational" that is at once sufficiently thick and sufficiently thin to plug into Smith's formula, and it turns out that there is no notion of rationality that is sufficiently thin for this purpose. Bear in mind that Smith's principle is supposed to be analytic. Even if we take the thinnest notion of rationality available, the one with the broadest appeal, and plug it into Smith's

formula, we still won't get a principle that's true by dint of meaning. To claim that we might approve of something that is wrong even if we were fully informed and fully logically consistent (i.e. fully thinly rational) is not *to contradict oneself*. Thus, no notion of rationality is thin enough to make Smith's principle true by dint of meaning. The search for a notion of rationality that is not too thick and not too thin is a project for those who would endorse a synthetic version of Smith's principle.

Like most Anglo-American philosophers, I'm a big fan of full rationality and full information. But let's not lose sight of the fact that I and others who share my proclivities are in the minority. "Wrong" simply does not mean "that of which we would disapprove if we were fully informed and fully rational." Of course, this doesn't mean that Smith's (modified) principle is false. For all we've said it might turn out that the things that are wrong are all and only those things of which we would disapprove if we were fully informed and fully rational. The point for now is simply that this principle can't be *analytic*. To deny it is not to contradict oneself.

2.4.3 The Specter of Relativism: Jackson, Smith (Again), and Lewis

Like Michael Smith, Frank Jackson (1998) defends a form of analytic naturalism, a view that goes under the headings "analytic descriptivism" and "moral functionalism." The essence of his view is this. Collect all of the standard ethical platitudes ("Lying is generally wrong," for example) and call that collection *folk*

morality. Call whatever folk morality would become after sufficient critical reflection *mature folk morality*. Then say that the property of wrongness is whatever property plays the wrongness role according to mature folk morality.[27] So it might turn out, for example, that if we were to submit our views to enough critical reflection we would all turn out to be utilitarians. If that turns out to be the case, then wrongness is—is *today*—the property of failing to maximize utility. Similar remarks apply to all moral properties.

This view is similar to Smith's in many ways. Both views are forms of analytic naturalism that make use of some improved form of our present moral thought to indirectly pick out the value-neutral properties on which the moral properties supervene. Not surprisingly, the same sorts of objections made against Smith's view in the previous section apply to Jackson's. Both views allow the nature of wrongness to be determined by the outcome of some kind of hypothetical rational discourse. As noted above, many people and many cultures have taken morality to have nothing to do with reason. Some have even claimed that acting rightly is a matter of putting reason aside in order to be guided by faith. Such people may be mistaken about the nature of morality, but we don't have to agree with their meta-ethical views to acknowledge that if they are mistaken, it's not because the they are incompetent users of words such as "wrong." If people with anti-rationalist meta-ethical views aren't merely confused about what "wrong" means, then we can't build a rationalist meta-ethics into the *meaning* of "wrong" without changing the subject. And yet that is what both

---

[27] Method follows Lewis (1970).

Jackson and Smith do. In other words they're both guilty of distorting the meaning of moral terms.

Here, however, I wish to focus on a different problem with Jackson's view (one that is, in a slightly different form, also shared by Smith's theory and Lewis' (1989) dispositional theory of value). The problem is that Jackson's account of the meanings of moral terms will render morality unacceptably relativistic. That is, it legislates against a kind of absolutism that many, if not most, moralists accept.

Jackson's moral functionalism allows the content of morality to be determined by accidental features of human history, thus failing to give the true moral principles the right "modal profile." A simpler version of Jackson's view would be one according to which we dropped the bit about critical reflection and said that wrongness is just whatever plays the wrongness role according to folk morality as it presently exists. The problem with this view is that it makes right and wrong a matter of what we happen to think right now. But that's not how morality works *according to folk moral theory*. (This highlights a point worth emphasizing. Folk moral theory includes its own meta-theory, a theory about its own nature.) Remember Sherlock Holmes and his crime-solving intuitions. The fact that we think something is wrong doesn't make it wrong any more than Sherlock Holmes' beliefs about the identity of the guilty party determine who committed the crime. Part of what we ordinarily think about morality is that the content of morality is not determined by what we happen to think about morality at present.

Jackson attempts to solve this problem by supplementing the simple view described above with his "critical reflection" clause, by putting the "mature" in "mature folk morality." But this doesn't really solve the problem. If we build into the idea of "critical reflection" that enough of it will take any moral system and turn it into the single true moral system, then we've accomplished nothing. All we've said is that wrongness is whatever would play the wrongness role in our theory once our theory has been perfected by a process that is guaranteed to perfect it. This is not, however, what Jackson has in mind. Jackson recognizes that, depending on where one starts, critical reflection might lead one to different places. Indeed, he acknowledges that his view allows for the possibility of relativism (Pg. 137). The problem is that his view actually goes a step further. Not only does it allow for relativism, it actually legislates against a rather commonplace, even dominant, form of absolutism. Most people, we'll suppose, subscribe to moral views that are not "absolutist" in the sense of being Draconian or rigid. Most people will occasionally say, "When in Rome…" and thus allow that different moral rules will apply to different people in different circumstances. We may grant all of this, but the fact remains that most people want to say that *some things*, sadistic torture for example, are wrong in an absolute, non-relative way. Jackson's view unwittingly rules out such beliefs, as the following example demonstrates.

Imagine a planet used as a penal colony where thousands of hardened criminals are deposited. These exiles have a code of behavior that plays the role of a moral system. They may even call it their "moral code" According to this

code individuals are allowed to kill and maim each other with little cause. Torturing fellow inmates is, on this planet, a common pastime. Certainly we would say that these people behave wrongly. Now forget the part about this planets' being a penal colony and assume instead that this is simply how their society developed on its own, or, better yet, how our society would have developed had things gone slightly differently in early human history. (Following Horgan and Timmons (1992), we can call this planet "Moral Twin Earth.") We still want to say that the way these people treat one another is horribly wrong. But what does it mean to say this? According to Jackson what we're saying is that these people are committing acts that have the property that plays the wrongness role according to mature folk moral theory. But whose mature folk theory, ours or theirs? (And remember, we can't assume that ours and theirs will end up being the same theory. If the maturation process is guaranteed to turn all folk theories into the one true moral theory then this mysterious maturation process does all the work and trivializes Jackson's view. Wrongness becomes whatever property plays the wrongness role according to the correct moral theory.)

There are two ways to go here, both of which lead to relativism and the rejection of commonplace absolutism. The first option is to say that wrongness varies from possible world to possible world, depending on what property plays the wrongness role in the improved version of the local folk morality. The second option is to is to "rigidify," to say that wrongness is the property that plays the wrongness role according to the improved version of folk morality in the *actual*

103

world, in our world, thus making wrongness the same property in all possible worlds.

Ask one of the folk if it's wrong to torture people for fun, and she'll say yes. Then ask her if it would still be wrong to torture people for fun if it were a socially accepted practice (like slavery was in past centuries). Presumably, she would answer that it would still be wrong. (It was upon realizing that this is so that we rejected the simpler version of Jackson's theory, the one without the maturity clause.) Next we'll suppose that this group of people who sadistically torture each other, the Twin Earthlings, are so far gone that no amount of critical reflection on their parts would lead them to change their minds. (Once again, we can't assume that this is impossible.) Would their torturous behavior still be wrong? Surely our folk moralist would say that even then it would still be wrong. Lesson: The folk aren't happy with the first option. The kind of behavior that we think of as morally abysmal is, according to the folk, still going to be wrong in worlds in which the inhabitants are so far gone that even their improved moral system looks wrong to us. The folk may go in for some amount of relativity, some adherence to "When in Rome…," but they won't stand for the kind of deep relativism that Jackson's theory requires, a relativism that rules out any kind of absolutist moral commitment.

What about the second option? "Rigidifying," unfortunately, just pushes the bump in the carpet. To rigidify in this case is to say that "wrong" refers to that property which would play the wrongness role according to an improved version of folk morality *as it exists on Earth at the present time.* This interpretation of

"wrong" builds into it a hidden indexical. It solves the problem of our having to admit that the Twin Earthlings way of life is not wrong, but it does so at a cost. We can say that their behavior is "wrong" only because we've stipulated that wrong *means* "wrong by improved Earth standards" as opposed to wrong in some more universal sense. The problem is that we don't ordinarily think of "wrong" as indexical, as making implicit reference to a particular place or time. Even if we sometimes apply moral terms with the understanding that they contain hidden indexicals, this need not *always* be the case.[28]

This point is illustrated by considering the consequences of the fact that Twin Earthlings will have their own use of the word "wrong." The Twin Earthlings say that torturing people for fun is "not wrong" and we say that it's "wrong." According to the rigidification story, there's no real disagreement here, just a miscommunication. The Twin Earthlings are simply saying that gratuitous torture fits in well with a matured version of *their* folk morality, and we're saying that gratuitous torture does not fit in well with a matured version of *our* folk morality. Such a disagreement would be like one between two people, one in New York and one in San Francisco, arguing over the phone about which city is "here." But we all know that that's not the right account of what's going on in this hypothetical disagreement. Ordinary moral thought supposes that there is what I call an

---

[28] Another way of putting this point, using terminology employed by Chalmers (1996) among others, is to say that by rigidifying we give terms like "wrong" an absolutist secondary intension and a relativist primary intension, but that the folk want an absolutist primary intension as well, at least in many cases.

ultimate asymmetry between those in the moral right and those in the moral wrong and, in this case, between ourselves and the Twin Earthlings. Yes, by rigidifying we can say that their behavior is ultimately and absolutely wrong in a non-relative way, but then they can say the same thing about us. And, yes, we can say, from our point of view, that their condemnation of our behavior is morally mistaken and that our condemnation of theirs is morally legitimate, but, once again, they can play the same game. At each level we can take refuge in our own point of view, but there's no denying that these two points of view are, as long as we insist that "wrong" and the like are indexical terms, ultimately symmetrical. And yet that is what folk moralists want to deny. Ordinary moral thought supposes that there must be a deep and ultimate moral asymmetry between those who engage in gratuitous torture and those who do not. Jackson's view unwittingly denies that this is so, implying that such behavior is either not necessarily wrong, or necessarily wrong but only from one particular point of view.

Whether or not we choose to rigidify, Jackson's moral functionalism cannot accommodate the fact that the folk take their morality to be *absolute* when it comes to certain things. Of course, ordinary moral thought could be wrong about this. (Indeed, I think it is.) The point, however, is not that this aspect of folk morality is correct, but rather that no satisfactory account of the meaning of moral terms should legislate against it. We don't want an account of moral language that renders ordinary moral thought incoherent. And yet that is what moral functionalism does by making right and wrong depend on (improved

versions of) what we happen to think here and now. Thus we find, in a curious twist, that we can't satisfy the demands of folk morality by allowing the content of true morality to be determined by the content of folk morality, or even by improved folk morality.[29]

Smith's view runs into similar trouble. Consider once again the morally depraved inhabitants of Moral Twin Earth. Let us suppose that their depravity runs so deep that they would still approve of the awful things they do if they were fully informed and fully rational. Is their behavior "wrong?" It's wrong, according to Smith, if we would disapprove of their behavior under conditions of full information and full rationality. The question then arises, "Who counts as 'we?'" If "we" includes the Moral Twin Earthlings then "we" would not disapprove of their behavior under those idealized conditions and so their behavior is not wrong. The most we can say if we're not willing to exclude them from the "we" is that their behavior is "wrong according to idealized Earth standards." If instead we restrict the "we" to Earthlings, then we simply build the relativism into the meaning of "wrong." Their behavior becomes absolutely wrong—they really do violate (improved) Earth standards—but their behavior is also "not wrong" in an equally legitimate, equally absolute sense, the sense associated with their own term "not wrong." And, as before, when we say that their sadistic ways are "wrong" and they say that these practices are "not wrong" we're not, according to this view, really disagreeing. Either way, Smith's view, like Jackson's, legislates

---

[29] As noted in Section 1.1, this argument applies to claims that quasi-realism would not have moral truths depend on our moral sensibilities.

against the possibility of making a coherent absolutist condemnation of the nasty Twin Earthlings.

In his contribution to "Dispositional Theories of Value" (1989) David Lewis offers an analysis of value, an analysis one might hope to parlay into an analysis of moral terms such as "right" and "wrong" and thus into a moral theory. He states that "something of the appropriate category is a value if and only if we would be disposed, under ideal conditions, to value it." Lewis identifies "valuing" with "desiring to desire," "ideal conditions" with "conditions of full imaginative acquaintance," and "we" with "the speaker and those somehow like him" where this "somehow" is used to produce as wide a "we" as circumstances will allow. Lewis is well aware of the problems identified in the foregoing discussions of Smith and Jackson and that these problems may threaten his own view. Nevertheless, he remains optimistic, but I will argue that such optimism is unfounded.

As before, everything turns on our interpretation of "we." Lewis outlines what he takes to be the strictest interpretation: "An *absolute* version of the dispositional theory says that 'we' refers to all mankind." Unless "mankind" refers to all possible individuals of whom moral evaluation is appropriate, I doubt that what Lewis calls the "absolute version" is absolute enough. Many would condemn the Moral Twin Earthlings for their nasty behavior and values regardless of whether they count as part of "mankind." We might suppose, as before, that Moral Twin Earthlings value sadistic torture and fail to value loving kindness, etc. Furthermore, we might assume that they are immune to

improvement through fuller imaginative acquaintance with these objects. If we want to avoid saying that loving kindness (or whatever) is not valuable, then, as before, we have to exclude the Twin Earthlings from the "we" (thus rigidifying). But can we do this while remaining true (or true enough) to the meanings of our words? Ordinary moralists want to say that the Twin Earthlings have mistaken values *and* that there is no ultimately symmetrical sense of "mistaken values" according to which this is not so. (At the very least, they want to say things like this *sometimes*.) Unless we are willing to say that ordinary moralists who stand by such assertions are not merely mistaken but actually *confused* about what it means for something to be a value, then we can't accept Lewis' analysis of "value" and any further analyses or theories that depend on it.

Note: In Section 1.1 I stipulated that moral realism requires a commitment to *mind-independent* moral facts. While one can interpret the views of Smith, Jackson, and Lewis as views according to which the moral facts depend on our minds, on our dispositions or sensibilities, I have chosen to classify these views as realist. I've done this primarily because the dependence of the moral facts on our dispositions/sensibilities is, according to these views, rather indirect. According to these views we could be dramatically wrong about the moral facts because the moral facts according to these views depend on *idealized* versions of our present dispositions/sensibilities and not on our actual

dispositions/sensibilities. Nevertheless, it is precisely this element of dependence, however indirect, that lands these views in trouble.

I also noted in Section 1.1 that some self-described moral realists such as John McDowell (1985) maintain that the moral facts are mind-dependent, and more directly so than according to Smith, Jackson, and Lewis. Thus, the arguments made in this section against Smith, Jackson, and Lewis—arguments aimed at the mind-dependence inherent in these views—apply to McDowell's "sensibility" theory (Darwall, Gibbard and Railton, 1992) and other like-minded views. More on McDowell and related views in Section 4.6.1.

2.4.4 Realist Relativism

One way to be an analytic naturalist is to be a realist relativist. A relativist is one who believes that moral claims, as they are ordinarily stated, can only be true relative to some moral framework and not true in some absolute sense. Moral truths are, according to this view, conventional, like truths about what is or is not legal. Is smoking marijuana legal? Yes, if by "legal" you mean "legal in Amsterdam,"; no, if you mean "legal in New York." Is it legal *simpliciter*? No such thing. All legality is legality relative to some system of laws. Likewise, says the relativist, for moral claims. The ancient Hittites had slaves and said that slavery is not wrong. We don't have slaves, and we say that it is wrong. Who's correct? We both are. Slavery is wrong$_{us}$, but it's not wrong$_{ancient\ Hittites}$. There is

110

no such thing as wrong *simpliciter*. All wrongness is wrongness relative to some

moral system, no one of which is privileged above the others.

Where does this view stand with respect to moral skepticism? There

certainly are facts about what is or is not wrong according to various moral

frameworks. If what it *means* for something to be wrong is for it to be wrong

according to the relevant moral framework, then it will be analytic that if

something has the value-neutral property of being wrong according to the

relevant framework, then it has the moral property of being wrong. (Here we're

once again making the *conservative* assumption that the candidate subvening

property, in this case the property of being wrong according to the relevant moral

framework, is value-neutral.) If this view of the meaning of moral terms is

correct, then moral relativism *and moral realism* are correct. (This is in spite of

the fact that most people tend to think of moral relativism as a kind of moral

skepticism.) The question, then, is whether or not this relativist account of the

meaning of moral terms is plausible.

I maintain that it is not, although there are, much to my amazement, others

who disagree.[30] When people at Human Rights Watch say that female genital

---

[30] Harman (Harman and Thomson, 1996) acknowledges that many people intend

their moral assertions to be absolutely true, but claims nonetheless that moral

relativism can stand as a form of realism. This is in spite of the fact that the

realist-relativist's account of the meanings of moral terms doesn't exactly fit with

what people have in mind when they use them. He points out that nearly

everyone prior to Einstein thought that there were absolute facts about what

objects are or are not in uniform straight-line motion. We now know that all such motion is relative to some spatio-temporal framework, but we don't say that there are no facts about what is or is not moving. Rather we say that there are such facts, but that these facts, contrary to what most people thought (and probably still think), are relative. Likewise, he argues, we shouldn't deny that there are moral facts and thus reject moral realism simply because it turns out that moral facts, like facts about motion, are relative, too.

This is a clever argument, but the physics-morality analogy ultimately fails to hold up. Suppose people were constantly zipping around near the speed of light so that different individuals typically resided in substantially different spatio-temporal frameworks. Under such conditions it would make little sense to speak of objects as simply "moving" or "not moving." And it would make even less sense if people were inclined to *fight* about which objects are or are not in motion. (Galileo and Bellarmine had such a fight, but such a fight would have been senseless had they both taken themselves to be talking past each other. Also, only uniform straight line motion is relative, and that was not the sort of motion in question.) Such circumstances are analogous to the present circumstances in ethics. The people of the world *do* have many different moral frameworks, and sometimes people are willing to fight each other over what they consider "right" and "wrong." In the physical case, we can get away with acting as if there really is an (absolute) fact of the matter about what is or is not moving because we all share the same spatio-temporal framework. The fact that motion turns out to be relative doesn't make much of a difference for the purposes to

mutilation is *wrong*, they aren't simply claiming that it's wrong according to some particular moral framework, the one to which they happen to adhere. It is not their intention to state a sociological fact about what is or is not allowed by the norms they and their kind endorse. Rather, they are saying that it's wrong in some *absolute* sense. They not only mean that it's wrong according to their moral framework, but also that their moral framework, unlike those of the perpetrators of female genital mutilation, is *the correct* moral framework, or close enough for human rights work. And they don't mean merely that their moral framework is correct *according to their moral framework*. They mean that it's correct *absolutely*. (Once again, it's a matter of *ultimate asymmetry*.) Likewise, when someone who is pro-female genital mutilation says that this practice is not wrong, he's not making a different sociological claim about his culture's mores, one that is, contrary to appearances, entirely compatible with everything the folks at Human Rights Watch have said. Rather, he's *disagreeing* with them. He's denying that what they've said is true.

While some would disagree, I am going to assume that the assertions made in the previous paragraph are correct and that the foregoing relativist account of the meanings of moral terms is inaccurate. What's more, I am going

---

which we ordinarily put physical language. In contrast, the way we talk about moral matters at present makes no sense on the assumption that moral claims can only be true relative to some moral framework and not absolutely. Witness the not-so-hypothetical discussion below concerning female genital mutilation. (More on the pragmatics of ordinary moral discourse in Chapter 4.)

to assume that it is *central* to what people have in mind when they make moral assertions that their assertions are absolutely true. In other words, I'm assuming that the foregoing account of the meanings of moral assertions is not "close enough" to serve as a set of reforming definitions for moral terms.[31] (More on reforming definitions in the next section.) I conclude that analytic naturalism in the form of realist relativism fails for lack of a plausible account of what people mean when they make moral claims. As most people recognize, moral relativism is really a kind of moral skepticism.

2.4.5 Other Forms of Analytic Naturalism

Once again, fundamental moral principles must connect value-neutral properties to moral properties, and such principles must be plausible as principles for which there is no further explanation. The strategy of the analytic naturalist is to look for fundamental moral principles among the analytic principles, a good idea given that analytic principles are generally self-explanatory, at least when they are properly understood. This strategy, however, is very ambitious. Fundamental moral principles that are supposed to be analytic must not only square with our first order moral intuitions (or so most ethicists assume); they must be true by dint of meaning. To deny such a principle, one must *contradict* oneself. Such principles are bound to be hard to find.

---

[31] See previous footnote.

In Smith's case we saw that a fundamental moral principle that equates moral behavior with idealized rational behavior, however appealing such a principle may be to us philosophy types, can't be analytic given the diversity of meta-ethical views conceptually competent people have held and continue to hold. In Jackson's case we saw that a moral principle that equates true morality with (improved) folk morality is itself inconsistent with the folk morality it attempts to accommodate, and Lewis' theory of value has similar problems. We also saw that analytic naturalism in the form of realist relativism depends on implausible assumptions about what people mean by their moral claims. Could other analytic principles fare better?   There's no way to say for sure, but there are good reasons to think that the answer is no.

Most, if not all, cultures have some sort of moral system and some set of evaluative terms that correspond roughly to our terms "right," "wrong," "good," "bad," etc. At the same time, these varied cultures (and subcultures) have radically different meta-ethical pictures. Some cultures have understood right and wrong in terms of what is or is not pleasing to the gods they worship. In contrast, some people understand right and wrong to be dictates of "pure practical reason." What's more, both the sophisticated rationalists of today and the simple pagans of the past are competent users of terms like "right" and "wrong." Given that both groups use these terms competently, one has to wonder what account of the meanings of such terms could be compatible with such radically different meta-ethical viewpoints and at the same time provide a foundation for substantive moral claims. Defining "right" and "wrong" in terms of

idealized rational behavior won't square with the pagans' use of these terms. Likewise, defining "right" and "wrong" in terms of that which pleases the gods won't suit us, rationalists included. It seems that the common denominator between our uses and their use of the term "wrong," for instance, is that the things that we call "wrong" and the things that they call "wrong" are, from each of our points of view, not to be done.[32] (This squares with the old non-cognitivist dictum that, contrary to appearances, the practical function of evaluative terms is to recommend, to guide action, but not to describe (Hare, 1952).) Moreover, we can't tie the meaning of terms like "wrong" to particular first order moral views since we can imagine cultures with first order moral views that are radically different from our own. We understand perfectly well what is being described when we suppose that, according to the Twin Earthling's moral code, gratuitous torture is not "wrong." It's hard to imagine how any analysis of moral terms could be compatible with such a wide range of first order and meta-ethical views without being trivial.

Perhaps the analytic naturalist will take refuge in the possibility of semantic indeterminacy for moral terms. The thought here is that it's not entirely clear what people mean by terms like "wrong" and that perhaps on some legitimate resolutions of such semantic indeterminacy there will turn out to be analytic moral principles that could serve as fundamental moral principles. Is it possible, for example, that while there are only a few people who *definitely* mean

---

[32] Of course, when people say that something is wrong they don't usually mean that it's wrong *merely* from a particular point of view.

by "wrong" "that of which we would disapprove if we were fully informed and fully rational," there are a great many others who *sort of* (i.e. indeterminately) think that this is what "wrong" means? Even this less ambitious claim about what people mean by "wrong" strikes me as implausible. How many people even *sort of* think it's *trivially true* that if we were fully informed and fully rational it would be impossible for us to approve of something that is morally wrong?[33] Likewise, how many people even *sort of* think that the maiming and torturing that goes on at Moral Twin Earth isn't wrong, or *sort of* think that a Moral Twin Earthling who says that torture is morally good isn't really disagreeing with us (because moral terms are implicitly indexical). It seems to me that the vast majority of people, past and present, have rather definite opinions about these things, and that most people view these commitments as *non-negotiable*. (Why this should be so will emerge in Chapter 3 in which we will consider the projectivist account of moral psychology. See Section 4.1.4 for an explicit discussion of the problem projectivist psychology poses for analytic naturalism.)

Remember that we're not talking about disagreements or uncertainty in ethical or meta-ethical opinion but rather about indeterminacy in the meanings of moral terms. Your average Joe may be willing to entertain the meta-ethical thesis that perfectly informed and perfectly rational versions of ourselves would be unable to approve of things that are immoral. But would your average Joe

---

[33]Note that semantic indeterminacy concerning the meaning of "rational" won't help here because, once again, there is no conception of rationality that is sufficiently thin to make Smith's principle true by dint of meaning.

say that this thesis might be *trivially true*, that it just follows from what people mean by "wrong" that this is so, that to think otherwise might be like believing in round squares? Who but very optimistic Anglo-American philosophers even *sort of* holds such views?

Of course, I can't prove that semantic indeterminacy won't save analytic naturalism, at least not without doing a great deal of historical and sociological research. But I'm sticking to my guns. Hanging one's hopes for moral realism on a favorable resolution of semantic indeterminacy in moral terms strikes me as a very desperate move. Rather than pushing ahead with this debate, however, I would rather table this issue until Section 4.7.2 in which I argue that this sort of moral realism (moral realism grounded in indirect moral principles) isn't worth trying to save in the first place.

For the analytic naturalist who's lost faith in the promise of semantic indeterminacy, there is yet another strategy to try. One might concede that no realist analysis of moral terms like "right" and "wrong" will be compatible with how these terms are actually used (even when we take into account semantic indeterminacy), but then suggest that we *change what we mean* by such terms. We might find an analysis that is close enough to what most people have had in mind (or sort of have in mind), and then stipulate that, from here on out, that's what we mean by "wrong" etc. For example, one might follow Smith's lead and say that henceforth "wrong" means "that of which we would disapprove if we were fully informed and fully rational." No one's to stop us from doing this. We're free to mean what we want by our words. The question, then, is whether or not it

makes sense to follow this course, and in order to make that determination we need to know what the alternatives are. In Chapter 4 I will argue that we should drop the language of moral realism altogether, and I will explain why this option makes more sense than the reforming analytic naturalist's option of preserving moral realism (in name only) by changing what we mean by our moral terms. In the meantime, suffice it to say that the analytic naturalist whose fundamental principles amount to reforming definitions of moral terms does not, strictly speaking, solve the skeptical problem. Whether or not this could be a good strategy nonetheless, is once again, a matter to which we will return in Section 4.7.2.

I must dispense with one piece of unfinished business before moving on. We said earlier that moral properties supervene on value-neutral properties, but in doing so we acknowledged that there is a chance that this might not actually be so. Once again, value-neutral properties are properties that we can ascribe to objects without logically committing ourselves to any value judgments. We said that low-level physical properties appear to be value-neutral, but that if it were possible to logically deduce evaluative claims from the ascription of some low-level physical properties then, strictly speaking, low-level physical properties would not be value-neutral, at least not all of them. In making that caveat, we were really talking about the possibility that some form of analytic naturalism might turn out to be true. We were talking about the possibility of true moral principles of the form "If something has such and such low-level, physical properties, then it has such and such moral properties," where such claims are

taken to be *analytic*. We've since made our case against analytic naturalism, and insofar as that case is sound, we can now say without qualification that moral properties supervene on value-neutral properties.

2.5 Synthetic Naturalism

2.5.1 The Synthetic Naturalist's Strategy

In the previous sections I explained why it's unlikely that the fundamental moral principles, the ones that forge the crucial links between the moral properties and the value-neutral properties on which they supervene, are going to be analytic. If such principles are not going to be analytic, they will have to be synthetic. Philosophers who embrace this conclusion while maintaining a commitment to naturalism are *synthetic naturalists*. The thought behind synthetic naturalism is not that "is wrong" and, say, "is an act that fails to maximize happiness" *mean* the same thing, but rather that two such terms may pick out the same property nonetheless, or pick out properties that are "necessarily equivalent."[34]

---

[34] For our purposes we can say that two properties that are necessarily coextensive are the same property. Whatever skeptical worries I raise about identities between moral and value-neutral properties can be reformulated as worries about analogous necessary equivalences.

There may be a variety of synthetic moral principles that are plausibly true so far as our first order intuitions are concerned. The problem with synthetic naturalism is not that it commits one to moral principles that are intuitively implausible. The problem with synthetic naturalism arises out of the fact that it, like all forms of moral realism, is implicitly committed to the existence of *fundamental* moral principles that are not plausible *as fundamental moral principles*. If synthetic naturalism is true, then there must be some synthetic moral principles that are *brutely true*, that admit of no further explanation. Thus, our skeptical problem with respect to synthetic naturalism is this: How could it be that some evaluative property (e.g. wrongness) is really the same property as some semantically distinct[35] value-neutral property (e.g. failure to maximize utility) without there being some story to tell about why those two properties are one and the same? Analytic naturalists deliberately sidestep this question by attempting to connect properties that are not semantically distinct. Synthetic naturalists typically fail to realize that this problem needs to be confronted. But it does need to be confronted, and synthetic naturalists will not have an adequate answer when the time comes to do so, or so I will argue.

Moore's "open question argument" sets the stage once again. In contrast to analytic naturalists like Smith and Jackson, the synthetic naturalist agrees with Moore that for any natural property it may well be an "open question" whether or

---

[35] A pair of "semantically distinct" properties is a pair of properties whose names are semantically distinct. As this term is used here, two properties can be semantically distinct and identical.

not something that has that property also has some evaluative/moral property, but maintains, contra Moore, that two such properties might nevertheless be the same. Consider someone who knows what "water" means and knows what it means for some substance to have the chemical structure H_2O but doesn't know that this is the chemical structure of water.[36] For such a person it's an "open question" whether or not these two properties have much to do with each other at all, but that doesn't mean that the property of being water and the property of being H_2O are not ultimately the same property. Likewise, say the synthetic naturalists, it might turn out that some natural property and, say, wrongness are really the same property even though, at this point at least, it's an open question whether or not they are in fact the same.

Synthetic naturalism avoids many of the pitfalls of other forms of moral realism. Unlike utilitarians and traditional deontologists, a synthetic naturalist need not commit herself to the truth of any direct and general moral principle and thus can avoid conflicts with first order moral intuitions. Unlike analytic naturalists, the synthetic naturalist avoids distorting the meanings of moral terms.

The strategy of synthetic naturalists is largely one of shifting the burden of proof onto the skeptic. They argue by analogy like so: Macro-economic facts supervene on lower-level facts. We haven't reduced macro-economic facts to lower-level facts, and perhaps we never will. Does this mean we should be

---

[36] The claim that water is H_2O is not brutely true, as it can be further explained. This claim is simply serving as an example of a synthetic claim that is plausibly taken to be necessarily true.

macro-economic skeptics? Surely not. Why, then, should the complicated and not-completely-understood supervenience of moral facts on lower-level facts drive us toward moral skepticism?[37]

To meet the synthetic naturalist's challenge, we skeptics have to explain what it is about the moral case that's special. We need to explain why we can accept non-reductive, supervenience relationships in many domains,[38] but not between moral properties and the lower-level properties on which they supervene.

2.5.2 Synthetic Naturalism and the Problem of Fundamental Moral Principles

The synthetic naturalist points out that many high-level properties supervene on lower-level properties in the absence of principles that reduce the former to the latter. Why, she asks, is this a problem for moral properties? The answer has to do with the infamous "is"/"ought" divide, or, in our terms, the divide between the evaluative and the value-neutral properties. All of the synthetic naturalist's models for how high-level properties supervene on low-level properties in the

---

[37] This sort of argument is found in Boyd (1988) and Sturgeon (1988).

[38] As it happens, I don't think that the relation between macro-economic facts (etc.) and lower-level facts is non-reductive, but my argument here will not depend on my taking this view. Rather, I will make my case based on considerations that are specific to the domain of value.

absence of reductive principles linking them (macro-economics, psychology, natural kinds, etc.) are all cases in which both the subvening and the supervening properties fall on the same side of the evaluative/value-neutral divide. But moral properties, as we know, are evaluative properties, and they are supposed to supervene on lower-level, value-neutral properties. This is what distinguishes moral supervenience from the other cases.

The synthetic naturalist claims that the relationship between wrongness and the lower-level properties on which it supervenes is like the relationship between water and H_2O. Scientists, using a combination of theory and observation, managed to uncover the essence of water, to find out what it really is (or so we'll assume). Likewise, philosophers (with a little help from their friends in other disciplines) may someday uncover the essence of wrongness, and they will do so not by *a priori* theorizing alone, but through a complex interplay between theory and observation.

It's a compelling story, but its being compelling depends on the fact that it's not yet finished. The problem with this story is that it can never have a satisfying ending. The reason is that in the final act we will inevitably encounter *fundamental* moral principles, principles that link evaluative properties such as wrongness to the value-neutral properties on which they supervene, and these principles, because they are fundamental, will admit of *no further explanation*.

Consider the following abridged and oversimplified history of our understanding of water and H_2O. We start out with some water. At some point, someone hypothesizes that water has the chemical structure H_2O. We perform a

number of experiments to test this hypothesis, and after a while we come to the conclusion that, indeed, it does have this structure. As this abridged history suggests, we were willing to entertain the idea that water is H_2O at the outset, but we didn't believe it until we heard more of the story, more about the specific nature of the relationship between being water and being H_2O, about how the fact that water is H_2O explains why water has the observable properties it has.

At the outset, we might look upon some moralist's hypothesis concerning what wrongness really is in the same way that we looked upon the hypothesis that water is H_2O before it was confirmed. That is, we might see it as an intriguing possibility and one about which we would like to hear more. The problem is that as the story of wrongness unfolds we're going to run into a "hypothesis" about which there is *no more to hear*. That is, at some point we're going to be confronted with a *fundamental* moral principle. It's going to be some principle of the form "Anything that has such and such value-neutral property is going to have such and such moral property," and there's going to be no explanation for why that's so.[39] The principle won't be analytic—recall that we're discussing *synthetic* naturalism—so it won't be self-explanatory in the way that analytic claims are. It will have to be brutely true.

To get a sense for how queer (Mackie, 1977) this is, let's make the case more concrete, first with a direct moral principle and then with an indirect one.

---

[39] I am not saying that no moral principles of this form can be explained. One must bear in mind that we are talking about fundamental moral principles of this form which, by hypothesis, cannot be explained.

Consider a utilitarian theory, one according to which an act is wrong if and only if it fails to maximize aggregate happiness. It doesn't seem crazy to think that, after some complex process of observation and theorizing, we might discover that the true essence of wrongness is failure to maximize aggregate happiness (or some such thing). However, our willingness to entertain this hypothesis depends on our thinking that there is some further story to tell about why this is so. It depends on our thinking that the aforementioned process of observation and theorizing will leave us with some kind of deeper understanding of why wrongness turned out to be—indeed *had to be*—failure to maximize aggregate happiness and not something else. But what if this utilitarian principle turned up at the *end* of the story rather than at the beginning? What if it were not a hypothesis to be fleshed out by further investigation, but rather the Holy Grail of moral theory, the Buck-Stopper, the Grand Finale? This would be most unsatisfying: "Wrongness just *is* failure to maximize utility? That's it? Nothing more to say?" We'd all want our money back.

Or would we? Perhaps it's just that we've yet to consider the right sort of synthetic fundamental moral principle. Like all direct and general moral theories that have been explored to any significant degree, old-fashioned utilitarianism isn't very plausible. But might we not be satisfied if we were presented with a different set of principles, ones with a little more flair and/or subtlety? (Ones that are indirect!) Might we not be satisfied with principles more like the ones proposed by Smith, Lewis, and Jackson, understood this time as synthetic? I doubt it. That's because all such principles, no matter how plausible at the

outset, will have to straddle the evaluative/value-neutral divide. Take, once again, the case of wrongness. Some value-neutral property will have to be identical to wrongness. You're going to look at this value-neutral property and you're going to think that, as far as you can tell, something that has this property may or may not turn out to be wrong. (Remember that the *synthetic* naturalist agrees with Moore that such questions will appear to be open.) And then the synthetic naturalist is going to *assert* that anything that has this property is wrong. And then you're going to ask "Why?" (After all, the question appeared open to you.) And then he's going to tell you that your question "Why?" has *no answer*. He's going to tell you that the principle connecting these two properties is just *brutely true*. And that's when you're going to want your money back.

Synthetic naturalism turns out to be something like a pyramid scheme. We're happy to leave Moore's question open with regard to some moral principles on the implicit assumption that out there somewhere there is some more complicated story to tell about what will close it. The problem is that no matter what story we tell about some moral principles, there will have to be other moral principles for which there is no story to tell about why they're true. Moreover, these fundamental moral principles will have to have a certain structure (value-neutral property on the left, moral property on the right), and any principle that has that structure is going to cry out for explanation. Synthetic naturalism can't fulfill its promise—not even *in principle*—because it will always land us back where we started, with a putative connection between a moral

property and a value-neutral property and a desire to understand why that connection holds.

One might object that we are holding synthetic naturalism to an unfairly high standard, that this is simply another round of unreasonable skepticism. After all, one might say, we shouldn't expect the process of moral inquiry to ever end. New answers always spawn new questions. Such is the mark of a healthy discipline. If synthetic naturalism's great failure is that it always leaves us asking more questions, what's so terrible about that?

The problem with synthetic naturalism is not that it will always leave us with unanswered questions, but rather that it will inevitably leave us with *unanswerable* questions which look like questions that *should have answers*. Consider the following two dialogues.

Suppose someone asks why the water that was in the pan on the deck disappeared. You explain that the sun's heat caused it to evaporate. And then that person might ask you why the Sun is hot. And you might say something about nuclear fusion taking place in the Sun's core. And then this person might then ask you why nuclear fusion takes place under those conditions. And in response you might, if you were so inclined, end up explaining the fundamental laws of physics as we understand them. And if that person asks you why the laws of physics are the way they are you would, at that point, have to throw up your hands and say, "I don't know why those laws are the actual laws. Perhaps your question has no answer. Maybe they just are the way they are." It would be

nice to be able to say more, but we all agree that the fact that at some point our physical explanations come to an end is no cause for skepticism about physics.

Now consider the moral version. You say that some act is wrong. Someone asks why. You say because it's a lie. She asks you why lying is wrong. You say that it's wrong because to lie is to treat someone as a means only. And then she asks why it's wrong to treat someone as a means only. And then you say that it's wrong because doing so is something of which we would disapprove if we were fully informed and fully rational. And then she asks why it's got to be wrong to do something of which we would disapprove if we were fully informed and fully rational. Isn't it *possible*, she asks, that someone who was fully informed and fully rational could disapprove of something that was not wrong? And then you simply say, "No. An act is wrong if it's something of which we would disapprove if we were fully informed and fully rational. Don't ask me why because there is no further reason. That's just the way it is."

The ending of the first dialogue is not entirely satisfying—one would like to know more—but at the same time it doesn't make one feel *cheated*. Whereas the ending of the second dialogue does make one feel cheated. (At least, it makes *me* feel cheated.) The fact that it ends with such an unsatisfying "because I said so" tends to undermine whatever merit there might have been in the answers that were given along the way. Of course, this dialogue could have played out in any number of ways, and perhaps the way I told it isn't the most plausible way. Still, the point remains that no matter how it's told the ending is going to be the same in terms of what's bothering us. All such dialogues have to

end with some kind of baldly asserted moral principle that leaves one asking, "So that's it?" Synthetic naturalism is implausible not because the explanations it can support have to come to an end, but rather because its explanations will inevitably come to a disappointing end.

2.5.3 Mathematics and Consciousness: Synthetic Naturalism's Partners in Innocence?

A standard response to any kind of local skepticism is to argue that it's not so local, that the arguments used to raise doubts about the present subject matter could just as easily be used to raise doubts about some other, presumably legitimate, subject matter. Such a response is open to the synthetic naturalist. Most notably, she might argue that the skeptical arguments made above could just as easily be used to raise doubts about the existence of mathematical facts and facts about consciousness.

Consider the claim that five plus seven equals twelve. We might wonder what makes this claim true. Some mathematical nominalists would argue that this claim states a highly general contingent fact and that it is made true by how the world happens to be. That however, is not the standard view. Most philosophers take mathematical truths to be true necessarily, and we shouldn't assume that they're wrong. Might this claim be analytic, true by dint of meaning? Again, that is not the standard view. Mathematical truths are typically taken to be

synthetic. So what makes this claim, which is plausibly taken to be synthetic and necessary, true? We could offer a proof, one that derives this claim from some set of premises. Such a proof might count as a good explanation of what makes this claim true, but in taking that approach we would end up back where we started. That is, we'd end up with a bunch of premises about which we might wonder what makes *them* true, and this would happen no matter how many proofs we gave. Ultimately, we'd end up with a bunch of axioms that are not analytic (or so the history of 20$^{\text{th}}$ Century analytic philosophy suggests), that are supposed to be necessarily true, and for which we have no further explanation concerning what makes them true. This situation is remarkably similar to the one which, according to the foregoing discussion, spelled doom for synthetic naturalism. Must we also conclude that mathematical facts are doomed? If so, what does this say about the argument given above against moral realism?

One way out, the way I favor, lies in a nominalist theory of mathematics, one according to which mathematical facts reduce to some kind of physical and/or mental facts. (I favor this option, however, not because it strengthens the case against moral realism, but rather because it seems to me correct on independent grounds.) Failing a nominalist treatment of mathematics, I would attempt to distinguish the mathematical case from the moral case via a modest appeal to Mackie's (1977) argument from relativity. The fact that people from markedly different cultural backgrounds agree so readily about which mathematical claims are and are not true suggests very strongly that there is a fact of the matter about which of them are or are not true. The fact that there is a

considerable degree of disagreement among moralists doesn't mean that there are no moral facts, but it does mean that "mathematical realism" has a point in its favor that moral realism does not. Their circumstances are not identical.

At the end of the day, however, I suggest that we judge moral synthetic naturalism on its own merits. I doubt that any synthetic naturalists would be inclined to give up on moral realism were they to become convinced that nominalism in the philosophy of mathematics was true. Nor, I suspect, would any moral skeptics change their minds about moral realism as the result of our reaching different conclusions concerning the foundations of mathematics. Insofar as that's the case, the alleged parallel with mathematics is a red herring. I don't find the fundamental principles that synthetic naturalism requires implausible simply because they are supposed to be synthetic, fundamental, and necessarily true. I find them implausible because they connect value-neutral and evaluative properties in the absence of any possible explanation for why those properties are so connected. Either you react to the possibility of such principles with skepticism or you don't, and whether or not you do probably has little to do with your views concerning the foundations of mathematics. It is, I think, a side issue at best.

Then there's the issue of consciousness. Mental states are supposed to supervene on physical states, which means that there must be principles that explain which brain states go with which mental states, "psycho-physical bridge laws." It's plausible enough that some mental states, especially those that fall under the heading of *qualia*, cannot be analyzed in terms of physical states. If

that's right, then the psycho-physical bridge laws must be synthetic. It's also plausible that the psycho-physical bridge laws are true *necessarily*. Furthermore, it's plausible that the psycho-physical bridge laws, some of them at least, will be *fundamental*. This means that there may well have to be a set of principles that are synthetic, necessary, and brutely true. Does the argument made above concerning the implausibility of synthetic, fundamental, and necessarily true moral principles speak equally well against the plausibility of facts about consciousness?

Here, too, I would emphasize certain differences between the two cases. I would begin, however, by acknowledging that, in a sense, the argument given above *does* speak in favor of skepticism about consciousness. Consciousness is, to my mind, *very mysterious*, and I doubt that I would believe in it were I not conscious myself![40] (Here I refer to phenomenal consciousness, that which poses the "hard problem" (Chalmers, 1995).) However, in the case of facts about consciousness there is an *even stronger* argument in their favor, namely the quasi-Cartesian argument to the effect that I cannot coherently doubt my own consciousness. While I'm left with no choice but to believe, it does strike me as strange that some brain states should give rise to some states of consciousness without there being any further explanation for why *those* brain states go with *those* states of consciousness. I like to think that there is *some* explanation for why the psycho-physical bridge laws are as they are, even it's one we humans are incapable of grasping, but if it turns out that there is no such explanation, I

---

[40] I confess that I'm not entirely sure what I mean when I say this.

don't see what choice I have other than to simply accept that that's the way it is. The fact that a good argument against the reality of consciousness can be countered by an even better argument doesn't mean that the same good argument can't be made in a different domain with greater success.

It is also worth noting that the psycho-physical bridge laws and the nature of the supervenience of the of the mental of the physical are amenable to empirical investigation in a fairly straightforward way.[41] There is no comparable research program underway in ethics. The supervenience of the mental on the physical is, in some cases, mysterious but undeniable. The supervenience of the moral on the value-neutral is just plain mysterious.

2.6 Non-Naturalism

So far we have considered and rejected analytic naturalism and synthetic naturalism. While this exhausts moral realism's naturalistic incarnations, we've not considered views that reject naturalism, for example, the view held by G. E. Moore (1959) according to which moral properties are metaphysically special non-natural properties. Similarly, I've made no serious attempt to take on traditional religious meta-ethical views according to which God plays an essential

---

[41] The study of the neural correlates of consciousness is booming. For recent attempts at neurologically based theories of consciousness see Damasio (1999) and Llinas (2001).

role in the determination of moral facts. This is in part because I find that I cannot take such views seriously myself, but mostly because I have little to offer in the way of convincing arguments to those whose inclinations are otherwise.

While my commitment to naturalism has played an important role in restricting my targets to views that are naturalistic, my commitment to naturalism has played virtually no role in the arguments that I've made against those views. Surely this comes as no surprise, but it is worth noting if only because some readers, upon on seeing where my arguments have led, may wish to reject them on the grounds that they depend on dubious metaphysical assumptions that have been smuggled in under the cover of naturalism. I do not believe that this is the case. My commitment to naturalism has, for the purposes of this essay, amounted to little more than a refusal to take seriously those views that invoke God, Moorean non-natural properties, Platonic forms, and the like.

2.7 Summary and Conclusion

Moral realists who remain unmoved by my arguments will likely respond to them in one of two ways. Some will make their rejoinders in the form of particular objections to particular steps in my arguments. Others, however, will claim that the entire project is misguided and that my arguments, which are fine so far as they go, demonstrate nothing more than that a certain metaphysically ambitious brand of moral realism is mistaken. Such critics might say something like this:

Ethics is not about discovering "moral facts" and "moral properties." It's not about finding out *how things are*, but about deciding what one *ought to do*. It is fundamentally and irreducibly *practical*, and if one insists on treating it as if it were a theoretical enterprise on a par with chemistry, one is bound to be disappointed. A life may be *described* from the point of view of natural science, but a life—that is, a *human life*—must be *lived* from within the *moral point of view.* To the self-described moral skeptic who refuses to acknowledge this basic fact of human existence, there can be no answer, and that is all your arguments show. From the point of view of science, morality is a wash, but the scientific point of view, useful though it may be, is not the only point of view. And—more to the point—it is not the point of view most pertinent to the moral enterprise.

Moral Realism with a capital "R" may be hopeless, but there is another kind of moral realism, equally deserving of the name, against which your arguments fall flat. It is an ethics of *reasons* not *facts*. When presented with a principle for action, the question is not, first and foremost, "Is it true?" but rather "Can I, as one rational agent among many, accept such a principle as a guide to action?" One answers such a question not by undertaking a search for *facts* that might *correspond* to such a principle, but by engaging in a process of *reflection*, one that employs certain *standards* for the evaluation of actions, standards which, though

not scientific, have authority for those who have adopted the moral point of view....[42]

While I have presented my arguments using words like "facts," "properties" and "supervenience,"—words that, to some, positively reek of "scientism"—the arguments I've presented cannot be brushed aside so easily. At the end of the day, all moral realists (sophisticated practical reason theorists, "Cornell realists," sensibility theorists, and virtue ethicists included) want to say that moral principles such as "Lying is wrong," (or more sophisticated variants thereof) are true. Nowhere have I insisted that such principles have to be made true by the existence of physical objects or, worse yet, metaphysical objects. I've only insisted that there be some story to tell about what makes such principles true. Moral realists are free to respond to this request by telling a story that characterizes right and wrong in terms of what sorts of reasons we have for acting and which of these reasons are capable of surviving certain kinds of rational reflection, and so on. (This is the approach taken by Smith.) When, in each case, such stories are presented the metaphysical question will still be there: *What makes that story true?* What makes it the case that *this* value-neutral property (say, the property of being unable to survive a certain kind of rational reflection) goes with *this* moral property (say, the property of being wrong). Is the claim that connects these properties supposed to be analytic? If

----

[42] The point of view expressed in this passage is similar to that of Korsgaard (1996).

so, there are problems. Synthetic? If so, there are problems. Conceiving of ethics as an investigation of "practical reason" rather than a search for "moral facts" is, so far as my arguments are concerned, simply a matter of stylistic preference.

I would like to emphasize the fact that my case against moral realism does not depend on a refusal to acknowledge the moral point of view. My question is not "Why should I care about being moral?" posed from the point of view of someone who shows no interest in doing so. It's not a matter of "What's in it for me?" a question to which the answer, even if moral realism is true, might be "nothing" or "nothing so far as your self-interest is concerned." Rather, my critique of moral realism is an *internal* one. "What makes my moral beliefs true?" is a legitimate question from within the moral point of view, just as "What makes my scientific beliefs true?" is a legitimate question from within the scientific point of view. "Correspondence to physical fact" is a plausible answer to the scientific version of this question. The moral version demands an answer as well, even if the same answer will not do. Ordinarily we simply assume, if only implicitly, that this basic question about the nature of morality has a satisfactory answer. I've offered reasons for thinking that, contrary to common sense, this question *cannot* have a satisfactory answer, and if I'm right then we have discovered a problem with moral realism *from within the moral point of view*. It's not as if we've refused to hear what moralists have to say about the nature of their enterprise and then—Lo and behold!—discovered that moral realism is false. One need not be a committed moral skeptic to think that there must be some story to tell about

what makes ordinary moral beliefs true, and one can conclude that there are deep problems with all such stories without restricting one's attention to stories about the nature of the physical world. Moral skepticism, I claim, is a natural reaction to a sincere attempt to make sense of morality on its own terms.

Let us take a moment to summarize the argument presented in this chapter. We concluded early on that moral realism requires the existence of true moral principles, claims that ascribe moral properties to things in virtue of their value-neutral properties. We also noted that, given the supervenience of the moral on the value-neutral, some of these principles must be necessarily true. Next, we concluded that some of these necessarily true moral principles must be brutely true or fundamental, since it's doubtful that we could explain the truth of the necessary moral principles as a whole in terms of contingent truths and/or non-moral necessary truths. We noted that all moral principles can be divided into general and specific moral principles and concluded that the fundamental moral principles cannot be specific. We noted that all moral principles can be divided into analytic and synthetic moral principles, as well as into direct and indirect moral principles. Moore's open question argument demonstrates that direct and general analytic moral principles are implausible. Next we examined some attempts at analytic naturalism, theories that rely on indirect and general analytic moral principles. We found that such views inevitably distorted the meanings of moral terms, often building into them unacceptably rationalistic or relativistic commitments. Finally we considered synthetic general moral principles, both

direct and indirect, and concluded that this approach, synthetic naturalism, inevitably leaves us with unanswerable questions that look like questions that should have answers. Thus we eliminated each of the possible types of fundamental, necessarily true moral principles, leaving us with the surprising conclusion that there are no true moral principles and that moral realism is false.

Regrettably, the above argument is not without loose ends. At many points its progress depends on my making bald assertions concerning what seems "reasonable," "plausible," "mysterious," etc. While many of these assertions are debatable, I don't think that any of them is crazy, or even terribly extreme. And yet somehow, when one puts them all together, one is left with the rather extreme conclusion that moral realism is false and that there is no fact of the matter concerning what's right and what's wrong. How did this come to pass?

The Dutch artist M. C. Escher is known for his captivating renderings of impossible objects, among them his *Waterfall* which depicts a downwardly zigzagging aqueduct that, through one of Escher's trademark visual twists, manages to deposit its load at its source. In focusing one's attention on any part of this image, one finds nothing to offend the eye. Every piece of stone, every splash of water, flows seamlessly into the next without the slightest hint of absurdity. It is only when one steps back to consider the image as a whole that its impossibility becomes apparent.

So, too, with moral realism. In life outside philosophy, we're given little incentive to step back from the totality of our first order moral beliefs so that we

140

might evaluate our implicit assumptions concerning the nature of moral truth. And, yet, it is only through such a process of stepping back that the problems with moral realism become apparent.[43] In this chapter I've attempted to bring these problems into focus, but these efforts, I imagine, will fail to convince many readers. Moral realism is, as I've said, very much a part of common sense, and common sense beliefs tend to fare well against technical and esoteric philosophical arguments. All the same, I will assume for the remainder of this essay that moral realism is false. My hope is that even readers who remain unconvinced by the arguments of this chapter will at the very least acknowledge that moral skepticism is a reasonable response to a challenging metaphysical puzzle and a position the consequences of which are worth exploring.

The aim of this chapter has been entirely negative, to build a case against moral realism. In the next chapter we will consider a positive account of the phenomenon of morality, the seeds of which were sown centuries ago by David Hume (1960, 1978) and that today draws support from a large and growing body of scientific evidence. This account will, first, help us understand why moral realism is so compelling in spite of the fact that it's false and, second, lay the groundwork for my case against revisionism, to be presented in Chapter 4.

---

[43] To step away from one's first order moral beliefs in this way is not to abandon the moral point of view, but rather to put certain aspects of the moral point of view to one side so that other aspects, those that we are ordinarily inclined to take for granted, may be thoughtfully considered.

Chapter 3

Moral Psychology and Projective Error

In the previous chapter we concluded, in spite of common sense, that moral realism is false. This raises an important question: *How is it that so many people are mistaken about the nature of morality?* We should be wary of any philosophical view according to which most people are profoundly mistaken about an important aspect of ordinary life. To become comfortable with the fact that moral realism is false we need to understand how moral realism can be so wrong but feel so right.

To that end, we turn to moral psychology, and in so doing we will lay the groundwork for Part II in which we'll address the practical questions that are the ultimate subject of this essay. In the previous chapter I made a case for anti-realism. Here I make a case for projectivism, or, at any rate, a picture of moral psychology that is broadly projectivist in flavor.

This picture has been evolving since Hume's time, but has gained widespread acceptance among experimental psychologists only in recent decades. In my opinion, the most complete and philosophically useful elaboration of this picture is in Jonathan Haidt's "social intuitionist model," described in detail in a paper entitled, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," (2001). Most of the empirical results discussed in this chapter are mentioned in Haidt's paper, and to avoid

redundancy I will in each of these cases cite Haidt's source rather than Haidt himself.

Sections 3.2, 3.3, 3.5, and 3.6 are drawn primarily from Haidt's discussion in which he defends his social intuitionist model against more traditional, rationalist theories of moral psychology. Haidt does not necessarily endorse the metaphysical picture for which I argue.   Nor does he argue for the revision of our moral thought and practices, although he does argue that common sense moral psychology is typically mistaken, a conclusion which tends toward a limited sort of revisionism at the very least. In Section 3.4 I build on the material presented in Sections 3.2 and 3.3 along with familiar ideas from cognitive science and evolutionary theory in an attempt to answer the main question with which this chapter is concerned: "Why does moral realism appear to be true even though it's false?" The material presented in Sections 3.2 and 3.3 is for the most part empirical, and I've chosen to present it in advance of the more speculative and theoretical material in Section 3.4 in order to provide a firmer footing for the latter discussion. Nevertheless, the argument advanced in these three sections is usefully summarized with a bit of rearranging: Why does moral realism appear to be true? Because moral experience has a perceptual phenomenology (Section 3.4). Why does it have a perceptual phenomenology? Because it is "easy" (Section 3.4), i.e. because it is something we do intuitively (Section 3.2) and, more specifically, emotionally (Section 3.3). Sections 3.5 and 3.6, concerned with moral development and the illusion of rationalist psychology respectively, provide crucial background information for subsequent chapters.

The case for anti-realism advanced in Chapter 2 was grounded in characteristically philosophical or metaphysical considerations rather than in the sorts of psychological and evolutionary considerations that will be discussed in this chapter. An injection of psychological/evolutionary considerations does, I believe, push people in the direction of anti-realism, but such considerations do not by themselves make for an adequate case against moral realism. What these scientific considerations suggest is that moral experience can be explained without appeal to moral facts, and this provides *some* people with adequate reason to give up on moral facts.[1] But many people's interest in morality is not primarily explanatory, and as a result there will naturally be those who, when confronted with evolutionary/psychological explanations of moral phenomena, say, "So what? That doesn't mean there's no fact of the matter about what's right or wrong." And they're correct. The evolutionary/psychological picture is highly suggestive of anti-realism, but arguments based on such considerations alone do not settle the meta-ethical question. They tell us what *is*, but they tell us nothing about the existence or absence of objective *oughts*, at least not in any straightforward way.

In the previous chapter I attempted to get at the heart of the meta-ethical matter by focusing on what I take to be the central meta-ethical question: *What could make moral claims true?* I came up empty handed after a rather laborious

---

[1] Harman (1977, 1996) argues that considerations such as these give us reason to give up on absolute/non-relative moral facts rather than moral facts *simpliciter*, but I maintain that this amounts to giving up on realism. (See Section 2.4.4)

search through the space of possible answers and concluded that realism is false. In a more reasonable world, the philosophical arguments given in the previous chapter would be enough to convince people that realism is false. In our world, however, the most forceful considerations are likely to come from elsewhere. I have in mind two sources, one relatively familiar and one new. The first and more familiar source is the emerging empirical picture of moral psychology that is the subject of this chapter. The second and unfamiliar source is a novel appreciation of the practical advantages of transparently anti-realist moral practices, to be discussed in Chapter 4.

3.1 Projectivism

Blackburn describes projectivism and identifies Hume (1960, 1978) as its canonical proponent:

> Let us call the Humean picture of the nature of morality, and the metaphysics of the issue, projectivism. On this view we have sentiments and other reactions caused by the natural features of things, and we 'gild or stain' the world by describing it as if it contained features answering to these sentiments, in the way that the niceness of an ice cream answers to the pleasure it gives us (1993a, pg. 152).

Suppose that we say we *project* an attitude or habit or other commitment which is not descriptive onto the world, when we speak and think as though there were a property of things which our sayings describe, which we can reason about, know about, be wrong about, and so on. Projecting is what Hume referred to when he talks of 'gilding and staining all natural objects with the colours borrowed from internal sentiment' or of the mind 'spreading itself on the world,' (1984, pg. 171).[2]

Projectivism is a cluster of closely related theses that comes in two main forms depending on the meta-ethical view to which it is attached. Realist projectivists maintain a form of subjectivist realism while non-cognitivists and error theorists favor projectivist anti-realism. The central tenet of projectivism is that the moral properties we find (or think we find) in things in the world (e.g. moral wrongness) are mind-dependent in a way that other properties, those that we've called "value-neutral" (e.g. solubility in water), are not. Whether or not something is soluble in water has nothing to do with human psychology. But, say projectivists, whether or not something is wrong (or "wrong") has everything to do with human psychology. More specifically, projectivists maintain that our judgments of right and wrong depend crucially on our emotional or affective responses (or, as

---

[2] See Hume (1878; I, iii, 14.)

Hume would say, on our "sentiments" or "passions").[3] Finally projectivists

maintain that our encounters with the moral world are, at the very least,

somewhat misleading. Projected properties tend to strike us as unprojected.

They appear to be really "out there," in a way that they, unlike typical value-

neutral properties, are not.

The projectivist's claim that moral phenomenology is misleading in this

way raises a metaphysical question about the existence of moral properties and

the truth of moral realism. Whether or not a certain kind of things exists depends

on the criteria a thing must meet in order to count as a thing of that kind. For

example, if being a vampire requires nothing more than being a person who

drinks blood then vampires almost certainly exist. Most of us would agree,

however, that vampirehood is more demanding and that therefore, given certain

reasonable assumptions, vampires do not exist. Likewise, in determining

whether or not moral properties are, like vampires, fictitious, we projectivists must

consider whether or not projected wrongness can be genuine wrongness, and

likewise for other moral properties. One might conclude that realism and

projectivism are compatible, a view I reject for roughly the same reasons I reject

the views of Smith, Jackson, and Lewis—too much mind-dependence. (More

about projectivism and realism in Section 3.4.) Projectivists who conclude that

realism and projectivism are incompatible face a further question, namely

---

[3] Other, non-moral forms of projectivism, most notably projectivism about color,

typically assert that the properties with which they are concerned are mind-

dependent, but not dependent on our emotional responses.

whether or not projectivism is compatible with our ordinary moral thought, discourse, and practices. Most people speak and think as if their moral judgments state facts about which things possess which moral properties. Is it a mistake to go on doing this once we've accepted projectivism and rejected realism? Blackburn (1984, 1993a) claims that it isn't, a claim he defends under the rubric of *quasi-realism*, a sophisticated offshoot of traditional non-cognitivism (Stevenson, 1944; Ayer, 1952; Hare, 1952; Gibbard, 1990). Richard Joyce (2001) adopts a similar position in his defense of *fictionalism*. In contrast, I recommend that we change the way we think and talk about morality, a position I call *revisionism*. I will assess the relative merits of these views and others in Chapter 4. For now let us simply note that accepting projectivism, in addition to committing one to the aforementioned psychological theses, forces one to confront both anti-realism and revisionism. Many projectivists have embraced anti-realism, but very few have embraced revisionism.

Projectivism predates the rise of psychology as an experimental science, but it is nonetheless a psychological thesis and one to be accepted or rejected on the basis of empirical research. I maintain that the existing research on moral psychology overwhelmingly favors the acceptance of projectivism in some form or other. I discuss this research in the next two sections.

3.2 Moral Intuition and Moral Reasoning

Until recently, moral psychology was dominated by rationalist models of moral judgment according to which moral judgments are primarily the products of reasoning and reflection (Haidt, 2001, pg. 816).[4] A large and growing body of research suggests that rationalist models are incorrect and that they should be replaced with what Jonathan Haidt calls a "social intuitionist model," one according to which moral judgments are primarily the products of immediate intuitive responses rather than episodes of reasoning:

> The central claim of the social intuitionist model is that moral judgment is caused by quick moral intuitions, and is followed (when needed) by slow, ex-post facto moral reasoning…. The words "intuition" and "reasoning" are intended to capture the contrast made by dozens of philosophers and psychologists between two kinds of cognition. The most important distinctions are that intuition occurs quickly, effortlessly, and automatically, such that the outcome but not the process is accessible to consciousness, while reasoning occurs more slowly, requires some effort, and involves at least some steps that are accessible to consciousness (Pp. 817-8).

---

[4] As we'll see later (Section 3.6), rationalist models of the kind proposed by earlier theorists and implicitly suggested by folk moral psychology take even "snap" moral judgments to be the products of reasoning.

I should emphasize that the above characterization of "reasoning," is not the only one available. One might adopt any number of more inclusive characterizations of reasoning, including ones according to which any mental process that produces a judgment (e.g. intuition, as described above) is an instance of reasoning. In the Western intellectual tradition, reasoning is generally held in high esteem, and people who wish to promote greater respect for intuitive thought may therefore wish to characterize intuition as a form of reasoning. While I am all in favor of giving intuition its due, I will nevertheless proceed using Haidt's terminology, thus drawing the relevant distinction as one between intuition and reasoning rather than between intuitive reasoning and some other form of reasoning. I do this in part because I believe that Haidt's terminology is appropriately similar to the terminology employed by Hume (1978; Book III, Part I) in his seminal discussion of reason and its role in moral thought.

The respective roles of intuition and reasoning are illuminated by considering people's reactions to the following story.

Julie and Mark are brother and sister. They are travelling together in France on summer vacation from college. One night they are staying alone in a cabin near the beach. They decided that it would be interesting and fun if they tried making love. At the very least it would be a new experience for each of them. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy making love but decide not to do it again. They keep that night as a special secret

between them, which makes them feel even closer to each other. What
do you think about that, was it OK for them to make love? (Haidt,
Bjorklund, and Murphy, 2000)

Haidt (2001, pg. 814) describes people's responses to this story as follows:

Most people who hear the above story immediately say that it was wrong
for the siblings to make love, and they then set about searching for
reasons. They point out the dangers of inbreeding, only to remember that
Julie and Mark used two forms of birth control. They next try to argue that
Julie and Mark could be hurt, even though the story makes it clear that no
harm befell them.[5] Eventually many people say something like "I don't
know, I can't explain it, I just know it's wrong."

This moral question is carefully designed to short-circuit the most common
reason people give for judging an action to be wrong, namely harm to self or
others, and in so doing it reveals something about moral psychology, at least as
it operates in cases such at these. People's moral judgments in response to the
above story tend to be forceful, immediate, and produced by an unconscious

---

[5] One might take issue with this claim, but it does nevertheless seem to be the
case that subjects' intuitions about the wrongness of Julie and Mark's actions are
more immediate and confident than whatever judgments subjects tend to make
about the possibility of harm.

process (intuition) rather than through the deliberate and effortful application of moral principles (reasoning). When asked to explain why they judged as they did, subjects typically gave reasons. Upon recognizing the flaws in those reasons, subjects typically stood by their judgments all the same, suggesting that the reasons they gave after the fact in support their judgments had little to do with the process that produced those judgments.

Under ordinary circumstances reasoning comes into play after the judgment has already been reached in order to find rational support for the preordained judgment. "When faced with a social demand for a verbal justification, one becomes a lawyer trying to build a case rather than a judge searching for the truth," (2001, pg. 814). Haidt notes that subjects who have condemned a victimless action as immoral often set out to find a victim, sometimes attempting to change the facts of the story in order to create one. For example, when asked about whether or not it is okay to clean a toilet with the national flag, many people who said that this action was wrong supported their judgment by claiming that this action might prove harmful to the agent. Such efforts are instances of a more general phenomenon that Haidt calls "moral dumbfounding," which is defined as "the stubborn and puzzled maintenance of a judgment in the absence of supporting reasons," (Haidt, Bjorklund, and Murphy, 2000). As Haidt points out, the very existence of moral dumbfounding indicates that, at least in some cases, the judgment comes first while the reasons for that judgment are made up afterwards.

Are these isolated phenomena? Perhaps people's responses to stories such as these tell us about the psychology of sexual taboo violation, but not about moral psychology in general. For all we've said here, it's possible that most moral judgments are caused by moral reasoning and that only moral judgments with respect to a few isolated topics are produced by intuition. There are, however, good reasons to think that this is not the case.

To begin, the picture of moral judgment as an intuitive process is part of a more general pattern observed by social psychologists. Haidt (2001, pg. 819) explains that, "The emerging view in social cognition is that most of our behaviors and judgments are in fact made automatically (i.e. without intention, effort, or awareness of process)."[6] Moreover, the claim that moral judgments are made primarily on the basis of moral intuition should ring true to students of contemporary ethics. Ethical theorists often compare their enterprise to the scientific enterprise of fitting theory to data, with the primary difference lying in the fact that for ethicists the "data" are moral intuitions rather than observed facts. This sort of moral theorizing is well exemplified by Judith Jarvis Thomson in her classic treatment of "The Trolley Problem"[7] (more on the Trolley Problem in Section 3.3) and described explicitly elsewhere by Thomson (1990) and by John Rawls (1971).

Linguist George Lakoff (1996) argues that the political differences between conservatives and liberals (differences over issues such as abortion,

---

[6] See also R. B. Zajonc (1984).

[7] See "The Trolley Problem" in Thomson (1986).

welfare reform, capital punishment, and affirmative action) are manifestations of deep differences in peoples' moral worldviews. According to Lakoff, the maintenance of different worldviews by liberals and conservatives allows them to "see" the world differently from one another such that what is obviously correct, reasonable, and "just common sense" to a conservative strikes a liberal as ludicrous, and vice versa. Lakoff points out that, from the point of view of cognitive science, nothing is "just common sense," and that when judgments are made on the basis of what we call "common sense" they are usually products of complex and *unconscious* cognitive processes (Pg. 4). He argues that moral judgments are more often formed as a result of metaphorical thought than as a result of reasoning and that getting someone to agree with you is more often a matter of getting her to "see things your way" than one of pointing out potential flaws in that person's (alleged) reasoning or assumptions. For example, to convince someone that we were right to have entered the Gulf War, your best bet would not be to reason about the costs and benefits of doing so or about abstract principles which might determine when one should or should not get involved in such conflicts, but rather to compare Saddam Hussein to Hitler, a madman who must be stopped. Likewise, if your aim is to breed opposition to American Gulf War policy, you would be wise to compare that conflict to the Vietnam War, an unacceptably bloody and costly endeavor (Haidt, 2001, pg. 825). Lakoff offers a complex cognitive model of liberal and conservative thought by which he attempts to make sense of the clusters of positions these two groups hold, the language they use to describe and defend their positions, and the difficulties

154

people on opposite sides of the political spectrum have in communicating with one another. The details of Lakoff's model, though fascinating, need not concern us here, however. The take-home point for our purposes is, once again, that moral and political judgments are, according to Lakoff, generally produced by intuition rather than reasoning.

Further support for the idea that moral judgment is largely intuitive comes from research in primatology. In summarizing the moral (or proto-moral) tendencies and capacities of non-human primates, Frans de Waal cites the prevalence of sympathy-related traits such as attachment, emotional contagion, special treatment of the disabled and injured, and cognitive empathy (the ability to trade places mentally with others); norm-related characteristics such as the following of prescriptive social-rules, internalization of rules, and anticipation of punishment; an understanding of reciprocity exhibited in concepts of giving, trading, and revenge as well as moralistic aggression against violators of reciprocity rules; and characteristics related to the maintenance of community life such as peace-making, the avoidance of conflicts, community concern, and negotiation (de Waal, 1996, pg. 211). Human morality appears to be an outgrowth of primate morality.[8] But primates, as far as anyone can tell, do not make moral (or proto-moral) judgments by means of effortful and deliberate reasoning involving a series of introspectively accessible steps. It's much closer to the truth (if not perfectly accurate) to say that sophisticated non-human primates such as chimpanzees produce their moral judgments intuitively. It

---

[8] See also Wright (1994).

seems unlikely that the psychological mechanisms at work in human morality should be radically different from those that play a strikingly similar role in the social lives of our closest relatives.

Reasoning in the sense defined above is not emotional.[9] With this in mind, the best argument in favor of embracing the social intuitionist model as a *general* model of moral psychology may come from research on the relationship between emotion and moral judgment, a relationship that is both very tight and very general.

3.3 Emotion and Moral Judgment

Recent work in psychology, cognitive science, and primatology suggests that moral judgments are, for the most part, intuitive judgments. But what kind of intuitive judgments are they? After all, we humans—well adapted creatures that we are—have intuitions about a wide variety of things, for example, linguistic intuitions (Pinker, 1994) and physical intuitions (McCloskey, 1983). Moral intuition shares a number of features with other kinds of intuition such as these, but it is different in at least one crucial respect. Our intuitive ability to determine whether or not a sentence is grammatical or where a falling object is likely to land

---

[9] Or, rather, when it is emotional it is only emotional incidentally. For any reasoning one might go through, one could go through that reasoning in an unemotional way.

does not appear to depend on our emotional capacities. Our moral intuitions, on the other hand, appear to depend crucially on our emotional capacities, or, to use Hume's term, on "sentiment." One might even go so far as to say that, as a general rule, moral intuitions *are* emotional responses. While that may be something of an overstatement, current research on emotion and its relation to moral and social behavior suggest that it may not be at all far from the truth.

Consider the famous case of Phineas Gage, a Nineteenth Century railroad worker who, as a result of an accidental explosion, passed a three-foot iron rod through the front of his brain (Damasio, 1994, pp. 3-9). Gage not only survived the accident; at the time he appeared to have emerged with all of his mental capacities intact. After a two-month recuperation period Gage was pronounced cured, but it was soon apparent that Gage's mind had been compromised by his accident. Before the accident he was admired by his colleagues for his industriousness and good character. After the accident, he became lawless. He wandered around, making trouble wherever he went, unable to hold down a steady job due to his anti-social behavior.

Doctors at the time did not understand why Gage's lesion, which was primarily to the ventromedial portion of his frontal lobes, had the profound but remarkably selective effect that it had. More recent cases of patients with similar lesions have shed light on Gage's injury.[10] Antonio Damasio (1994, pp. 34-51)

---

[10] The advent of brain imaging has provided further converging evidence for the hypothesis, suggested long ago by Gage's case, that the prefrontal cortex plays a special role in the regulation of moral and social behavior. Raine *et al* (2000)

reports on a patient named "Elliot" who suffered a brain tumor in roughly the same region that was destroyed in Gage's case. Like Gage, Elliot maintained his ability to speak and reason about topics such as politics and economics. He scored above average on standard intelligence tests, including some designed to detect frontal lobe damage, and responded normally to standard tests of personality. However, his behavior, like Gage's, was not unaffected by his condition. While Elliot did not develop anti-social tendencies to the extent that Gage did, he, too, exhibited certain peculiar deficits, particularly in the social domain. A simple laboratory probe helped reveal the subtle but dramatic nature of Elliot's deficits. When shown pictures of gory accidents or people about to drown in floods, Elliot reported having no emotional response but commented that he knew that he used to have strong emotional responses to such things. Intrigued by these reports, Damasio and his colleagues employed a series of tests designed to assess the effects of Elliot's damage on his decision-making skills. They asked him, for example, whether or not he would steal if he needed money and to explain why or why not. His answers were like those of other people, citing the usual reasons for why one shouldn't commit such crimes (Pg. 46). Damasio and Saver followed up this test with a series of five tests of

---

found that individuals with antisocial personality disorder have, on average, eleven percent less prefrontal gray matter than control subjects. In an earlier study using positron emission tomography (PET; a brain-imaging technique) Raine *et al* (1994) found that murderers, as compared to normal subjects, exhibited decreased metabolic activity in the prefrontal cortex.

moral/social judgment (Saver and Damasio, 1991). As before, Elliot performed normally or above average in each case. It became clear that Elliot's explicit knowledge of social and moral conventions was as good or better than most people's, and yet his personal life, like Gage's, had deteriorated rapidly as a result of his condition (although he did not seem to mind). Damasio attributes Gage's real-life failures not to his inability to reason (in the sense above) but to his inability to integrate emotional responses into his practical judgments. "To know, but not to feel," says Damasio (Pg. 45), is the essence of Elliot's predicament.

The cases of Elliot and Gage are complementary. In Gage we see a connection between ventromedial damage and anti-social behavior. In Elliot we see a connection between ventromedial damage and emotional deficits (which almost certainly were present in Gage as well), resulting in poor prudential and social decision-making. Both patients are without "intellectual" deficits, suggesting a dissociation between "higher cognition" on the one hand and social, moral, and emotional capacities on the other. Taken together, these cases suggest that moral behavior (and social behavior in general) depends crucially on emotional capacities, a claim that would surprise many philosophers and (until recently) psychologists, but not David Hume.

Other cases provide dramatic support for this Humean picture of moral psychology. In a study that included Elliot as well four other patients with similar brain damage and similar problems with social behavior, Damasio and his colleagues found among these subjects a consistent failure to exhibit typical

159

electrodermal responses (a standard indication of emotional arousal) when presented with socially significant stimuli, though these patients responded normally to non-social, emotionally arousing stimuli (Damasio, Tranel, and Damasio, 1990). While the subjects in this study exhibit "sociopathic behavior" as a result of their injuries, they are not "psychopaths." A more recent study (Anderson *et al*., 1999) concerns two subjects whose ventromedial cortices were damaged at an early age (three months and fifteen months). These two subjects do exhibit characteristically psychopathic behavior—lying, stealing, violence, and lack of remorse after committing such violations—presumably because they, unlike the subjects with late-onset damage, did not have the advantage of a lifetime of conditioning via normal emotional responses.[11]

The study of "ordinary" psychopaths has proved illuminating as well. According to Blair *et al*., "The clinical and empirical picture of a psychopathic individual is of someone who has some form of emotional deficit," (Blair *et al*., 1997). Psychopaths exhibit a lower level of tonic electrodermal activity and show weaker electrodermal responses to emotionally significant stimuli than normal individuals (Hare and Quinn, 1971). A more recent study (Blair *et al*., 1997) compares the electrodermal responses of psychopaths to a control group of criminals who, like the psychopaths, were serving life sentences for murder or

---

[11] The early onset patients, unlike the late onset patients, showed deficits in their explicit knowledge of social and moral norms. Anderson and colleagues suggest that this deficit may occur due to an ability to experience feelings necessary for the normal acquisition of explicit knowledge of moral and social norms.

manslaughter. While the psychopaths were like the other criminals in their responses to threatening stimuli (e.g. an image of a shark's open mouth) and neutral stimuli (e.g. an image of a book), they showed significantly reduced responses to distress cues (e.g. an image of a crying child's face) relative to the control criminals, a fact consistent with the observation that psychopathic individuals appear to have a diminished capacity for emotional empathy.

Blair (1995) hypothesized that psychopaths' diminished capacity for emotional empathy should prevent them from drawing a distinction between what he calls "moral" and "conventional" rules and that psychopaths, as compared to other criminals, should make fewer references to the pain or discomfort of victims in explaining why certain harmful actions are unacceptable. Both of these predictions were confirmed. "Moral" transgressions were defined as those having negative consequences for the "rights and welfare of others" and included instances of one child hitting another and a child smashing a piano. "Conventional" transgressions were defined as "violations of the behavioral uniformities that structure social interactions within social systems" and included instances of a boy wearing a skirt and a child who leaves the classroom without permission. While the "normal" subjects (non-psychopathic incarcerated criminals) drew a general distinction between moral and conventional transgressions, the psychopaths did not. Normal subjects found a greater difference in permissibility and seriousness between moral and conventional transgressions than did the psychopaths. The most striking finding, however, concerned the psychopath's judgments concerning "modifiability." Each of the

transgression stories were set in school, and in each case the subjects were asked whether or not it would be permissible for the child to perform the transgressive action if the teacher had said earlier that such actions were permitted. Non-psychopathic criminals tended to say that the conventional transgressions would become permissible if the teacher were to explicitly allow their performance but that the moral transgressions would not be permissible in either case. Psychopaths, however, treated all transgressions as impermissible regardless of what the teacher says.[12] In addition, the psychopaths were, as predicted, less likely to appeal to the pain and discomfort of victims and more likely to appeal to the violation of rules in explaining why various transgressions are impermissible.

In attempting to assess the role of reason vis-à-vis emotion in the production of moral judgment and behavior, psychopaths provide crucial information. If moral reasoning is primarily responsible for the production of moral judgment and behavior, then we would expect those who exhibit diminished reasoning capacities to be the ones who make the most "moral

---

[12] While Blair predicted that the psychopaths would fail to draw the moral/conventional distinction, he predicted more specifically that the psychopaths would treat both types of transgression as normal subjects treat conventional ones. Nevertheless, he found that the psychopaths treated both types of transgressions as normal subjects treat moral ones. He attributes this result to the fact that his psychopathic subjects (all incarcerated) have an interest in demonstrating that they have "learned the rules."

mistakes." If, on the other hand, moral judgment and behavior are primarily the products of emotional response, then we would expect those with diminished emotional capacities (of the relevant kind) to exhibit the most morally abysmal behavior. While there can be no doubt that normal moral life is shaped by both reason and emotion, the study of society's most recalcitrant moral offenders points to a very clear answer.

Blair's psychopaths, Damasio's prefrontal patients, and the legendary Phineas Gage, each social disasters in spite of their intact reasoning abilities, vividly illustrate Hume's (1978; Book II, Part III, Section iii) infamous claim that, "'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger," Likewise, their apparent emotional deficits lend equal support to Hume's positive claim that the edifice of human morality rests on a foundation of "sentiment."

Recent brain-imaging research conducted by myself and my collaborators (Greene *et al*. 2001) has shed some light on the role emotion plays in the moral judgment of normal adults. After participating in a number of discussions of Judith Jarvis Thomson's classic paper, "Trolley Problem," (1986), I noticed a puzzling pattern in people's intuitive moral judgments concerning some of the hypothetical dilemmas discussed in Thomson's paper. In one such dilemma a runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save these people is to turn the trolley onto an alternate set of tracks where it will run over and kill one person instead of five. The question is whether or not it is morally acceptable to turn the trolley in order

163

to save five people at the expense of one? Most people say immediately that it

is. In another dilemma, a runaway trolley threatens to kill five people as before,

but this time the agent is standing next to a large stranger on a footbridge

running over the tracks, in between the oncoming trolley and the five people.

The only way to save the five people is to push this stranger off the bridge and

onto the tracks below. He will die as a result of doing this, but his body will stop

the trolley from reaching the others. Here the question is whether or not it's

morally acceptable to save the five people by pushing this stranger to his death.

Most people say immediately that it is not.

Thomson considers these dilemmas and the intuitions they elicit as part of

a puzzle for moral theorists: What makes it morally acceptable to sacrifice one

life to save five in the original case but not in the footbridge case? Answering

this question has proved surprisingly difficult. Many plausible answers have

been proposed, but they all run into trouble sooner or later,[13] which means that

---

[13] The following example taken from Thomson's discussion illustrates the point.

Kant (1993) says that one ought never treat someone as a "means only." With

this in mind, it might occur to you that what distinguishes these two cases is that

in the footbridge case one literally *uses* someone to save the others, whereas in

the original case one doesn't *use* the other person; he just happens to be in the

way. Consider, then, a case just like the original one only this time the track with

the one person on it loops around so that without that person there to block the

trolley it would come back around and kill the five people you were trying to save.

In this variation on the original case you must *use* the other person to stop the

the trolley problem, if it has a solution, has no obvious solution. The fact that it

has no obvious solution is interesting from a psychological point of view. This is

because nearly everyone judges immediately that it's acceptable to act in the

original case and not in the footbridge case in spite of the fact that it's very hard

to say why this is so (echoes of Haidt's "moral dumbfounding").[14] How do people

consistently manage to immediately give the "right" answers when there is no

obvious line of reasoning that would lead one to those answers?

In light of the above discussion, a natural hypothesis presents itself, as

well as a means of testing it. In the most general terms, one might suspect that

the uniformity in people's judgments concerning these cases is explained by a

uniformity in their patterns of intuitive emotional response. I made the following

more specific hypothesis: Pushing someone to his death with one's bare hands is

more *personal* than hitting a switch that will have similar consequences.[15] The

---

train. Nevertheless, it seems to most people that it's still okay to turn the trolley

in this new case, thus demonstrating the inadequacy of this (admittedly crude)

Kantian principle. While more sophisticated principles have been proposed, as

far as I know, no one has found a set of principles that adequately captures

people's intuitions concerning these and other cases.

[14] There is a parallel here with Chomsky's observation that individuals, even

young children, can adeptly apply the rules of grammar in the absence explicit

knowledge of those rules. See Pinker (1994).

[15] At least in this case. One can imagine cases in which hitting a switch would

cause someone's death in a way that is intuitively rather personal.

thought of harming someone in this personal way is likely to be more emotionally

salient than the thought of harming someone in a relatively impersonal way.

Because people have a robust, negative emotional response to the personal

violation proposed in the footbridge case they immediately say that it's wrong,

much as did Haidt's subjects in response to the taboo violations he proposed. At

the same time, people fail to have a strong negative emotional response to the

relatively impersonal violation proposed in the original trolley case, and therefore

revert to the most obvious moral principle, "minimize harm," which in turn leads

them to say that the action in the original case is permissible.

From individuals like Phineas Gage and Elliot and from a number of

neuroimaging studies,[16] we've learned that certain parts of the brain are

associated with emotional response. Brain scanning techniques such as fMRI

(functional Magnetic Resonance Imaging) are capable of detecting real-time

changes in the level of activity in various regions of the brain. Therefore, we

would predict that subjects passing judgment on dilemmas like the footbridge

case while undergoing brain scanning would show increased activity in at least

some emotion-related brain areas as compared to those passing judgment on

dilemmas like the original trolley case. We would also make the following

prediction concerning an interaction between people's responses and reaction

times. According to the above hypothesis, when people contemplate the

performance of a personal violation such as pushing someone in front of a

---

[16] For two useful reviews see Reiman (1997), Maddock (1999), and Davidson

and Irwin (1999).

trolley, they experience an emotional response which inclines them to judge that action inappropriate. Some people, however, *do* judge such actions to be appropriate. Unless such individuals are psychopaths, we would expect that when they give such answers they do so *in spite of* their emotions. We would predict, then, that such individuals would, on average, take longer to give their judgments than those who judge such actions to be inappropriate and longer than those passing judgments on dilemmas that do not elicit strong emotional responses.

Both of these predictions (neurological and behavioral) proved accurate. Three brain areas previously associated with emotional response (including one of the areas affected by Gage's accident)[17] were more active during contemplation of "personal" moral dilemmas such as the footbridge dilemma than during contemplation of "impersonal" dilemmas such as the original trolley dilemma.[18] We also found that reaction times for trials in which the subject

---

[17] The first of these regions lies in the medial prefrontal cortex, which is behind the forehead along the space that separates the two cerebral hemispheres (Brodmann's Area 9/10). (This area was one of the areas affected by Gage's accident.) The second is the posterior cingulate region, which is also along the space between the hemispheres, but closer to the back of the brain (Brodmann's Area 31). The third region is centered around the angular gyrus which is on the side of the brain towards the back (Brodmann's Area 39).

[18] Dilemmas were classified as "personal" or "impersonal" by independent coders according to the following criteria. Dilemmas concerning actions that are likely to

judged a personal violation to be appropriate were on average significantly longer than those of other types of trials. By implicating emotional response as a key determinant of moral judgment in normal adults, these data lend further support to the social intuitionist model. They also suggest a modification. While it may be true in general that emotions tend to drive moral judgment, moral reasoning may hold sway in some types of moral judgment. Constructing a more detailed psychological typology of moral judgments is a matter for further research[19] and theorizing, but the personal/impersonal distinction is, I believe, a step in the right direction.

Haidt aptly summarizes the Humean shift in our understanding of emotion and reason in moral psychology:

Journalists studying Congress sometimes say that to understand the legislative process one should "follow the money," and not the political rhetoric. Similarly, psychologists studying moral behavior should "follow

cause (a) serious bodily harm (b) to a particular person (i.e. someone who is "on stage" in the dilemma) (c) where this harm is not the result of deflecting an existing threat onto a different party are classified as "personal." All other moral dilemmas were classified as "impersonal." (Coders were given more extensive and specific guidelines than those described here.)

[19] This theory makes a straightforwardly testable prediction concerning psychopaths and patients like Elliot. We would expect them, unlike most people, to respond to the trolley dilemma and the footbridge dilemma in similar ways.

the emotions," and not be fooled by *ex post facto* moral reasoning.

Reasoning ability may be more salient, and tools for measuring moral

reasoning may be more fully developed, but reasoning may be little more

than the tail wagged by the dog. The dog itself appears to be moral

intuitions and emotions (Haidt, 2000).

3.4 Projective Error

Moral judgment is, for the most part, intuitive and emotional. So far so Humean.

But this alone does not amount to *projectivism* in the fullest sense. Consider the

projection metaphor. We patrons of the cinema experience a kind of illusion. Of

course, *we* know that what we see before us are mere projections and not real

people and places, but our visual systems are fooled, often to great effect. The

moral projectivist claims that in our encounters with the moral world we ourselves

are fooled (often to great effect). We "project" moral qualities onto the world and

consequently see it as if it possessed these qualities all by itself, no help from us.

Why does the world appear this way? Why do the things that we think of as

wrong strike us as positively *oozing* wrongness? Why do the things that we think

of as right appear to positively *glow* with rightness? Why don't we simply see the

objects of our moral interest as things that, though neutral in themselves, cause

us to have certain kinds of emotional reactions in response to them? Why are

we unaware of our tendency to project?

Our answers to these questions are far from complete. Many philosophers would say that we don't know why we have any phenomenological experiences at all (Chalmers, 1996), let alone why we experience particular aspects of the world as we do rather than some other way. Nevertheless, while we do not yet know why things seem as they do, experimental psychology has demonstrated that very often things are not as they seem. The mind is far from transparent to itself and often quite misleading about its own nature and the nature of the world. At the same time, evolutionary theory and cognitive science have helped us understand why it makes sense that we should fall prey to certain kinds of illusions and not others. Here we look to these disciplines to ease the pain of rejecting moral realism by explaining how an illusion of moral realism fits in well with our evolving understanding of the human mind and the forces that have shaped it.

There's a saying among cognitive scientists: "Easy things are hard," (Minsky, 1985, pg. 29). That is, the most computationally complex things we do are most often the things we do with the least effort and the least understanding of how we do them. For example, people have no trouble at all determining whether or not a given human face is male or female. We don't know *how* we do this. We just do it, quickly and effortlessly. Except in rare instances, we have no experience of *figuring out* whether or not the human faces we encounter are male or female even though these judgments are in fact extremely complex (Wiskott *et al*., 1995). From our point of view, human faces just *look male* or *look*

female. The complexity of these judgments is buried deep in the subconscious processes that make these judgments on our behalf.

In contrast to such "easy" tasks as sorting faces by sex, we find tasks such as memorizing the state capitals or finding the square root of 2,345 considerably more difficult. But these tasks, which require from us a relatively large amount of effort, are actually much simpler than the aforementioned face-sorting task in terms of their computational demands. They can be performed very quickly by very simple computers running very simple programs.

Why are so many of the things we find easy actually quite hard and vice versa? Part of the answer is that the forces of natural selection that have shaped our brains have lavished more R & D on some of our abilities and less on others. Sorting conspecifics by sex is a longstanding challenge to sexually reproducing species, and it's no surprise that our abilities in this regard reflect many millions of years of practice. Having the names of the state capitals at one's disposal and an ability to calculate square roots in one's head can be very useful under the right circumstances, but our use for such skills is a relatively recent development, and our lack of practice is revealed in our bumbling efforts to accomplish these tasks.[20]

---

[20] It's worth nothing that some people are more bumbling than others, and in different ways. Consider autistic savants who lack the most basic of social skills, who lack "common sense," but who can perform inordinately complex mathematical calculations in their heads (Dehaene, 1997). The existence of these individuals gives us insight into the sorts of tradeoffs that mother nature

What about moral judgments? Are they hard or easy? According to the social intuitionist model, moral judgments are, for the most part, easy—that is, easy from our point of view, not from a computational point of view.   It may sound strange to say that moral judgments are easy. After all, aren't moral judgments paradigmatically difficult judgments? Yes and no, but more no than yes. Most of morality is "common sense," and most of our moral judgments are made effortlessly and without any awareness of how they were formed. Consider, once again, the footbridge case. Nearly everyone says immediately and confidently that it's wrong to push the man in front of the trolley to save the other five people. While justifying this judgment in its broader context is very hard, arriving at this judgment is easy. It's a simple matter of intuition.

Of course, some moral judgments are hard. Choosing death for one's child is an obviously horrible thing to do, as is allowing both of one's children to die. A forced choice between these options, as in the dilemma made famous by the film *Sophie's Choice*, is fascinating and horrifying because it's a rare instance of a choice in which both alternatives are deeply morally repugnant. Naturally, moral dilemmas of this kind receive much attention, but we shouldn't lose sight of the fact that most moral judgments—the ones we don't bother to talk about, the ones that are just "common sense"—are not hard at all. Moreover, even these hard cases may be better characterized as easy cases gone awry. By this I mean that they are not hard in the way that difficult mathematical calculations are

---

has made in designing our minds. With the right kind of cognitive apparatus, hard can be easy and easy can be hard.

hard. They aren't cases in which one blindly pokes around looking for a clue or laboriously plods toward an obscure solution. Cases like Sophie's choice are not so much cases in which it's hard to see the right answer as they are cases in which it's hard *not* to see *two* right answers, or more accurately, two wrong answers. They are cases of double-vision, not blindness.

One may be inclined to say that moral judgments are hard because *moral reasoning* is hard, but, as noted above, moral reasoning and moral judgment are two different things. If the social intuitionist model is correct, moral reasoning is a largely *ad hoc* phenomenon, typically coming into play after the relevant moral judgment has already been rendered (Haidt, 2001, pg. 820). Recall the case of Mark and Julie, the two college-aged siblings who decided to have sex. When asked about their decision to have sex, most people immediately and confidently say that it was wrong for them to do so (easy). When asked to provide a reasoned justification for this judgment, people typically engage in "moral dumbfounding," slowly and laboriously producing half-baked justifications that they themselves often find lacking (hard). Once again, philosophers and lay moralizers alike say immediately and confidently that it's okay to turn the trolley away from the five people and onto the one person but not okay to push the one person in front of the trolley to prevent it from killing the five (easy and easy). Philosophers are still trying to produce a satisfying principled justification for this set of judgments that doesn't run into similar trouble elsewhere (hard).

Discussions of trolley type cases, like most philosophical discussions of normative ethics, implicitly—and in many cases explicitly (Thomson, 1990)

—suppose that moral intuitions do and should play the role of "data" in moral theorizing. That is, our intuitive moral judgments about particular cases are what's *given*, the phenomena for which moral theories must account. As a description of a practice, this jibes with the social intuitionist account. As Haidt says, intuitive moral judgments come first; the reasoned justifications for those judgments come later. Of course, moral reasoning is not always *ad hoc*.[21]

---

[21] Earlier I mentioned the possibility that the cases focused on by Haidt and others who reject rationalism in moral psychology are the cases most likely to elicit intuitive judgments and least likely to elicit reasoned responses. As noted earlier, there is good reason to think that moral judgment is intuitive in general, and the aforementioned results concerning psychopaths and brain-damaged individuals lend further support to this conclusion. Nevertheless, there may be more of a place for rationalist moral psychology than Haidt supposes. The contrast between the original trolley case and the footbridge case illustrates this point. The brain imaging data concerning these and other moral dilemmas suggest that cases such as the original trolley case do not elicit much of an emotional response. As noted earlier, my interpretation of people's tendency to judge the action proposed in the original trolley case to be morally acceptable is that, in the absence of an emotional response either way, people apply a simple consequentialist principle and thus opt for saving the greater number of lives. It seems to me that some judgments such as this one, those that I've called "impersonal," may be accurately described by a rationalist moral psychology. See Greene *et al.* (2001).

Occasionally philosophers have drawn (and acted upon) surprising moral conclusions about particular practices on the basis of reasoning from general moral principles. Some have claimed, for example, that we ought to drastically change the way we treat animals (Singer, 1975) or that it's wrong to spend one's money on middle-class luxuries while there are needy people in the world who could benefit immensely from those resources (Unger, 1996). But these esoteric moral judgments are clearly not the norm, and in a sense they are the "exceptions that prove the rule."   Such arguments are advanced in an attempt to overturn the conventional moral wisdom, to get us to think and act in spite of our ordinary, more facile moral judgments.[22]

Why does it matter whether moral judgments are hard or easy? It matters because easy psychological processes and hard psychological processes differ in their histories and characteristic phenomenologies. More specifically, by understanding moral judgment as merely one among many deceptively easy psychological processes, we can begin to understand why moral phenomenology should lead us into *projective error*.

The things that we do best, the things that feel easiest to us, are the things for which we have inherited special skills from our ancestors. These complex

---

[22] One might have the following worry: "What if these radical ethicists are right? What if we've come to many wrong moral conclusions about basic moral issues? How, then, can we say that moral judgment is easy?" Bear in mind that by "easy" I don't mean "easy to get right," but rather "produced by an effortless, intuitive process." Moral judgments can be easy in this sense even if they're all incorrect.

cognitive processes that nature has made easy for us are often associated with a *perceptual phenomenology*. Once again, when we look at a male face we don't *figure out* that it's male. It just *looks male*. And when we see that same face the next day from a different angle and in different lighting conditions we don't *figure out* that it's the same face. It just *looks familiar*. When we hear a sound coming from the left we don't *figure out* that it's coming from the left. It just sounds like there's something *over there*.

In contrast to the above cases of perception, we don't look at Magic Johnson and see him as HIV positive. Nor do we look at stars that are receding and see them as receding. This is because mother nature has not equipped us with the tools necessary for determining intuitively whether someone is HIV positive or whether a given star is moving away from us. We might have been designed so that individuals who are HIV positive just *look* HIV positive to us or so that stars that are receding just *look* like they're receding, but we weren't. As a result, making such judgments is hard, and our making them does not feel like a matter of simple perception.

I've been comparing our intuitive moral judgments to our intuitive judgments concerning other things such as the sex and familiarity of human faces. But if the anti-realism defended in the last chapter is correct, there is an important difference between our moral intuitions and these other sorts of intuitive judgments. Some humans really are male, female. Some faces really are familiar, and some auditory stimuli really are emanating from *over there*. I claim, however, that nothing is really right or really wrong. It's not hard to see

176

why mother nature would have endowed us with an intuitive ability to ascertain certain facts about our environment, but why would she endow us with an ability to ascertain non-existent moral facts?

Alas, mother nature is not as kind as one might hope, epistemically speaking. She doesn't care whether or not we have true beliefs. Or, more charitably, she only cares about the truth of our beliefs to the extent that our having true beliefs contributes to our reproductive fitness. That said, having true beliefs is generally advantageous in this regard. If, for example, there is a hungry tiger in the vicinity, one is, *ceteris paribus*, well-served by a belief that this is so, and likewise with any number of other beliefs about the state of the world. Moreover, we have developed special capacities for forming true beliefs about particular aspects of our environments that are especially important for our survival, for example beliefs about which of the objects we encounter are potential mates, threats, sources of food, etc.

We members of *Homo sapiens* are a highly social species. Most of the distinctive biological strategies we've developed depend crucially on our complex interactions with one another. Furthermore, we are a highly intelligent species, capable of serving our ends in any number of ways, including ways that impose great costs on others. As a result there are obvious benefits to humans having $moral_2$[23] sensibilities, i.e. to our being concerned to respect the interests of

---

[23] Recall that "$moral_2$" refers to morality *qua* concern for the interests of others, i.e. to those (all-important) aspects of morality that remain in the absence of moral realism (i.e. morality$_1$).

individuals other than ourselves. In the most general terms, what each of us

gains from not living with individuals who disregard the interests of those around

them far outweighs what each of us loses by showing consideration for the

interests of those around us (Axelrod, 1984; Wright, 2000). While there is an

ongoing debate about how morality$_2$ managed to evolve,[24] there is little doubt

among evolutionary theorists that morality$_2$ has evolved and that it has its

advantages. Its advantages, however, are importantly different from those

associated with our purely perceptual capabilities. As noted above, our ability to

tell at a glance who is male or female is advantageous because it allows us to

accurately represent mind-independent facts about our environment. In contrast,

our collective dispositions to help and refrain from harming one another are not

advantageous because they allow us to know something about how the world is

in and of itself, but because they help us survive as individuals in a social

context.[25]

So goes one version of the standard evolutionary account of the

emergence of morality$_2$.[26] Certain aspects of the story as I've told it may be

questioned, but I am going to assume that it's generally correct because my task

here is not to explain the existence of morality$_2$ but rather to explain the appeal of

morality$_1$. The existence of morality$_2$ is by no means irrelevant to this task,

---

[24] See, for example, Dawkins (1976) and Sober and Wilson (1998).

[25] Or as conduits for genes (Dawkins, 1976).

[26] I've told this story with a bit of an anti-realist gloss that some friends of

evolutionary moral psychology may not endorse.

however. This is *because an efficiently implemented morality$_2$ is likely to result in an illusion of morality$_1$*. By an efficiently implemented morality$_2$ I mean one that is *easy*, one that works more like our ability to sort faces by sex (quick, effortless, reliable) and less like our ability to memorize state capitals (slow, effortful, unreliable). As noted above, cognitive processes that are easy are likely to result in a perceptual phenomenology. Knowing that moral judgments are, in general, *easy*, it should come as no surprise that we *see the world in moral colors*, that we are in the unconscious habit of "gilding and staining" the world with moral properties that, according to the results of the last chapter, it simply doesn't have.[27]

---

[27]This projectivist account of moral phenomenology explains why we are inclined toward realism when it comes to "easy" moral judgments, but what about hard cases? Why do people often think that there is a fact of the matter about what's right or what's wrong, even when they don't know what it is?

First, there is the aforementioned "double vision" phenomenon associated with difficult moral dilemmas such as the one from *Sophie's Choice*. Here we still project moral properties onto the world and are thus inclined toward realism. The difficulty in such cases lies in too much projection, not too little.

I believe that most hard cases are cases in which two intuitive emotional responses are pitted against each other or in which an intuitive emotional response is pitted against a less impassioned concern for greater consequences as, for example, in the dilemma concerning whether or not to negotiate with terrorists. (Our emotions tell us to do whatever we can to save the real, live

individual(s) being held hostage—especially if we feel a personal connection to him/her/them—but our consequentialist tendencies warn against negotiating under such circumstances, as doing so may make things worse in the long run by encouraging future terrorist activities.) In cases such as these there is at least one intuitive emotional response at work, and that is, according to my theory, enough to promote, or at least sustain, an illusion of realism in the face of uncertainty.

Second, we might consider realism in the face of difficult moral dilemmas that elicit no intuitive emotional responses. I don't know if there are any such cases, but even if there are, the fact that so many moral dilemmas do appear to have clear right answers may lead people to conclude that there must be right answers even in cases in which the answers aren't clear. This is especially plausible for people whose moral beliefs are grounded in religious beliefs according to which there is a metaphysical moral order that makes sense even if it often fails to make sense to us. ("The Lord works in mysterious ways.")

Emotivist non-cognitivists typically look to projectivism to provide an account of the *meaning* of moral judgments. In doing this, they commit themselves to the rather implausible thesis that all moral judgments, however unconfident or otherwise emotionally detached, are expressions of emotional projections. Error theorists such as myself, however, can avoid this problem. I appeal to people's general tendency to project to explain people's general tendency toward realism, and this approach does not require that there be an episode of projection behind every realist moral judgment.

Something similar occurs in the domain of sexuality. There, as in the moral domain, we and other animals see the world through a metaphysically misleading, evaluative lens. Presumably, when a sexually receptive male baboon views a sexually receptive female baboon, he will likely see her as being very *sexy*, and her him. What's more, to a lustful baboon, the sexiness of the object of his/her lust presumably strikes him/her as *intrinsic sexiness*. This is not to suggest that baboons draw an explicit distinction between intrinsic and extrinsic properties and then judge that sexiness is a property of the former sort. The point is rather that, as far as we can surmise, baboons experience sexiness as being just as "out there" as anything else. A baboon has no inkling that the sexiness s/he's experiencing in any way depends for its existence on his/her own psychological make-up. But, of course, that other baboon's sexiness has everything to do with its admirer's psychological make-up. We know this because we whose minds are wired-up somewhat differently from those of lustful baboons fail to see baboons as sexy at all. (At least most of us don't.) Sexiness is in—or at any rate, depends crucially on—the mind of the beholder. Appropriately situated animals see each other as sexy not because there's an independently existing property, sexiness, that these animals can successfully detect, but because it serves such animals well to see each other as sexy. By seeing each other in a sexy light they can avoid the hassle of having to *figure out* that it would be advantageous for them to have sex with each other. And if this sort of a system lulls some of its more sophisticated users into false metaphysical

beliefs about the ontology of sexiness, so be it. Providing the world's creatures with true metaphysical beliefs was never the point.

Whether or not a tendency to see certain things as sexy lulls those who have it into projective error, there are certainly other domains in which such errors are committed. Any number of visual illusions illustrate this point. Here's one. A red spot appears on a screen followed almost immediately by an identical red spot slightly to the left. What one sees (or "sees") is the first dot *move* from right to left, and in so doing one sees a red streak connecting the two red dots. Of course, there is no red streak out in the world, though it certainly looks as if there is. What, then, is this red streak? Where is it? Is it *real*? The most natural thing to say about this red streak is that it is *projected*. It's not really there, but our minds put it there because it makes sense to put it there. It makes sense because, under ordinary circumstances, a sensation of an object followed by a sensation of a very similar object that is slightly displaced in space and time are almost certainly the sensory effects of a single moving object. Thus, it serves our interests to see such phenomena as single moving objects rather than as pairs of objects disappearing and appearing. By seeing things this way we don't have to *figure out* what's in front of us. Instead, our well-honed visual systems make such judgments for us. However, in letting our visual systems make things easy for us by turning what might have been a conscious inference into an immediate perception (or "perception") we run the risk of being fooled, of seeing something that isn't really there.

These sorts of projective illusions can be rather high level and extend into the social domain (Scholl and Tremoulet, 2000). For example, when people are shown computer displays in which geometrical figures move in certain suggestive patterns those people immediately and without prompting attribute desires (e.g. wanting to catch one of the other figures), emotions (e.g. frustration, anger), and personality traits (e.g. shyness, being a bully) to those figures.[28] (This effect is remarkably consistent across cultures and age-groups, starting as young as three years old.[29] There is some evidence suggesting that such attributions are made by infants and chimpanzees.[30] The point, of course, is not that the subjects in these experiments actually *believe* that the little squares and triangles moving around the computer screen have desires, emotions, and personality traits (although some of the younger subjects may in fact believe this). Rather, the point is that people can't help but *see* these displays in such terms:

The importance of such phenomena stems partially from the fact that these interpretations seem to be largely perceptual in nature—to be fairly

---

[28] In one case, for example, a square moves so as to appear to be chasing a small triangle, which in turn is being guided and protected by a circle. See (Heider and Simmel, 1944).

[29] See Morris and Peng (1994), Rime *et al.* (1985), Hashimoto (1966), and Berry and Springer (1993). Above cited in Scholl and Tremoulet (2000).

[30] See Gergely *et al*. (1995) and Cisbra *et al.* (1999). Above cited in Scholl and Tremoulet (2000).

fast, automatic, irresistible, and highly stimulus driven—despite the fact that they involve impressions typically associated with higher-level cognitive processing (Scholl and Tremoulet, pg. 299).

The way our minds are set up, we just see the world in social terms, and in some cases we can't help but project social properties onto parts of the world that in fact have no such properties. Of course, some objects in the world really do have desires, emotions, and personality traits, and our tendency to project those properties onto those objects involves no error.[31] Our capacity for social perception, like our capacity for the perception of moving objects, only gets us in trouble *some* of the time.

But I claim, once again, that our moral psychology gets us in trouble *whenever* we make moral$_1$ judgments. This distinguishes moral$_1$ judgments from all the other types of judgment to which I have compared them. Let us take a brief inventory of the relevant similarities and differences among these various kinds of judgment.

---

[31] Or so we ordinarily think. One might take the radical view that many cases of ordinary perception are really covert Gettier cases (Gettier, 1963). The thought is that the cognitive shortcuts we use actually prevent us from having genuine knowledge of the world. We get things right, not because we're lucky (as in the standard Gettier cases), but because natural selection has contrived to make us get things right, i.e. to project properties onto the things that happen to really have those properties.

Like judgments about the sex of human faces, the sexiness of baboons/people, the motion of moving objects, and the mental states of objects exhibiting certain kinds of motion, moral judgments are generally easy (phenomenologically) and are generally associated with a perceptual phenomenology. Moreover, except in the case of simple motion perception, these judgments have a distinctively social function. These are the most relevant similarities. The differences are more subtle. Because some individuals really are male/female, some objects really are moving, [32] and some of those moving objects really have the mental states we attribute to them, projectivist anti-realism with respect to these domains is a mistake. The most anti-realistic thing we can say about these cases is that our minds are so constructed as to lead us into projective error with respect to these things on certain occasions. I stress that while projection itself may be ubiquitous in these domains, it is only on certain occasions that such projections amount to projective *errors*. In the ordinary case one will project, say, maleness or motion onto something that happens to really be male or moving. Of course "happens to be" is a bit misleading. Mother nature has spent millions of years making sure this coincidence occurs.

Once again, the case of sexiness is importantly different from the cases of motion and importantly similar to the moral case because there are mind-independent facts about what's moving or not while there are no mind-independent facts about what's sexy or morally right/wrong. There may, however, be an important difference between the case of sexiness and the moral

---

[32] That is, moving, relative to the relevant spatio-temporal framework.

case in that a lack of mind-independent facts about what's sexy may not imply that there are no facts at all about what's sexy ("sexiness anti-realism"). That depends on what we're committed to in thinking that certain things are sexy. Let us once again consider what we are committed to in the moral case.

In Section 1.1. I stipulated that moral facts can't be mind-dependent in a straightforward way. In Chapter 2 I argued against forms of moral realism that would have the moral facts be mind-dependent in a less straightforward way, and those arguments work *a fortiori* against views according to which, contra my stipulation, the moral facts are straightforwardly mind-dependent. The central idea behind those arguments is this: To say of the nasty Twin Earthlings that what they do (gratuitously torture one another, for example) is only wrong from our point of view[33] is to give up on the idea that what they do is *really* wrong. (It is for this reason that "relativism" is a popular term for moral skepticism.) When we say that something is morally wrong, we mean that its wrongness is such that it would still be wrong even if our moral sensibilities had been different. That's why projectivism and anti-realism go hand in hand in the moral domain. Projectivism suggests that moral properties, at least the ones we think we're perceiving, depend on our moral sensibilities in precisely the way that ordinary moral thought does not allow.[34]

---

[33] Or only wrong from an improved version of our point of view, or only wrong because "wrong" is indexical. See discussion in Section 2.4.3.

[34] Contrary to what I've said here, many philosophers argue that that moral properties are, to use Mark Johnston's (1989) term, "response dependent," and

An analogous alliance does not necessarily exist in the domain of sexiness. If I say that so-and-so is sexy and you deny that this is so, I might just say, "Well, I find so-and-so sexy, but you don't, and that's all there is to it." In the moral case we might "agree to disagree" for various reasons, but insofar as we are moral realists we'll be sure to leave out the "and that's all there is to it." It's not clear that saying "and that's all there is to it," in the above exchange concerning so-and-so's sexiness commits one to an anti-realism about sexiness because it's not clear that the mind-independence of sexiness is essential to realism in that domain. Perhaps it is, but insofar as this isn't obviously the case, projectivist realism with respect to sexiness is a live option.

Once again, moral projectivism and moral anti-realism go hand in hand,[35] and one can start with either one in hopes of securing the other. As noted earlier, my strategy has been to reject moral realism on independent grounds and then, with an assumption of anti-realism already in the background, explain

that this is compatible with realism. See discussion of Smith, Lewis, and Jackson in Section 2.4.3..

[35] That is not to say that they strictly imply one another. Moral anti-realism would be true even if we failed to project, and realism would be compatible with projectivism if there were some sort of supernatural basis for moral truth such as God's will or Moorean non-natural moral properties. More realistically (i.e. naturalistically), it's possible that a qualified sort of realism along the lines of Blackburn's quasi-realism would be acceptable if certain contingent facts were otherwise. See Section 4.5

the apparent truth of moral realism by appeal to our psychological tendency to project. Others have employed the reverse strategy, beginning with a positive projectivist account of morality$_2$ and then arguing from there against realism on the grounds that projectivism alone can account for all of the relevant phenomena, thus relieving us of any reasons we might have had for believing in genuine, unprojected moral properties (Harman, 1977). Once again, I favor the first strategy over the second because I take the problem to be fundamentally metaphysical, not epistemological—a matter of "How could things be that way?" rather than "What reasons do we have for believing that things are that way?" Either way, we may conclude that the widespread error embodied in moral realism is an error that makes sense—a mistake, one might say, that we were born to make.

3.5 The Shaping of Morality

Moral judgment is, for the most part, intuitive. Where do these intuitions come from? Are they innate, or are they learned? Haidt explains:

> Perhaps because moral norms vary by culture, class, and historical era, psychologists have generally assumed that morality is learned in childhood, and they have set out to discover how morality gets from outside the child to inside. The social intuitionist model takes a very

different view. It says that morality, like language, is a major evolutionary

adaptation for an intensely social life, built into multiple regions of the brain

and body, which is better described as emergent than as learned, yet

which requires input and shaping from a particular culture. Moral intuitions

are therefore both innate and enculturated (2000, pg. 827).

An examination of the moral, or proto-moral, lives of other primates suggests that

human morality has a powerful innate component. As noted in Section 3.2, non-

human primates exhibit sympathy-related traits such as attachment, emotional

contagion, special treatment of the disabled and injured, and cognitive empathy

(the ability to trade places mentally with others); norm-related characteristics

such as the following of prescriptive social rules, internalization of rules, and

anticipation of punishment; an understanding of reciprocity exhibited in concepts

of giving, trading, and revenge as well as moralistic aggression against violators

of reciprocity rules; and characteristics related to the maintenance of community

life such as peace-making, the avoidance of conflicts, community concern, and

negotiation (de Waal, 1996, pg. 211). Among the most striking of the above

chimpanzee traits is their capacity for what is hard to describe as anything other

than moralistic aggression. Haidt (2000) summarizes an incident originally

reported by de Waal (1982).

> [The] zoo-keepers had created the rule that no chimp would be fed until all
>
> members of the colony had moved from the outdoor island to the indoor

sleeping enclosure. The chimps themselves worked to ensure that

stragglers obeyed the rule, in part by showing hostility to latecomers. One

evening two adolescent females stayed out for hours, delaying the feeding

of the whole colony. The zoo-keepers finally isolated the two errant

adolescents, fearing for their safety if they faced the angry colony. The

next day, when they were released into the company of the colony, the

other members vented their hostility on them. In other words, the

members of the colony were aware that specific individuals were the

cause of the delayed feeding, they remembered this information overnight,

and they inflicted punishment the next day.

It is worth noting that, as far as anyone can tell, these chimps did not consciously

adopt this method for social regulation in the way that we humans have, say,

adopted various methods of voting. This sort of behavior clearly has a strong

biological basis, and insofar as we, one of the chimpanzee's closest biological

relatives, exhibit similar behavior, parsimony suggests that such behavior in

ourselves is continuous with theirs, that is, largely instinctive and not a rationally

designed cultural invention (Haidt, 2001, pg. 827).

Further empirical support for morality's strong innate component comes

from the various pathological cases we have already considered. Psychopaths,

who on average have reduced prefrontal gray matter (Raine *et al.* 2000), have a

hard time sorting social violations into what Blair (1995) calls the "conventional"

(e.g. exiting the classroom without permission) and the "moral" (e.g. pulling

someone's hair). (Not to mention psychopaths' more obvious moral abnormalities.) It would be strange if their biological deficits were not the primary cause of their abnormal behavior.[36] And, once again, certain patients who have sustained ventromedial damage at an early age, like psychopaths, show not only a deficit in moral behavior but a deficit in their explicit knowledge of moral norms. The fact that such selective deficits are possible suggests that we do not learn moral norms by means of a generalized learning process. We appear to have a special capacity for the acquisition of moral norms, a capacity that is dramatically compromised in certain pathological individuals.

Morality, however, cannot be completely innate, as moral norms and intuitions vary considerably across cultures. As noted above, the shaping of morality appears in many ways to be like the shaping of language. We humans have a special capacity for acquiring human language, that is, language of the specific form that is unique to and universal among humans (Pinker, 1994). Our having this capacity, however, is not sufficient for language acquisition. No one will ever learn English or Chinese without being exposed to those languages. Nevertheless, the ability to speak one's native language arises out of powerful instincts and is not acquired in the same way that other skills are acquired. To begin, there seems to be a critical period for language acquisition. Adults can learn a second language, but doing so is a difficult process that requires a great

---

[36] And it seems unlikely that this sort of biological deficit (eleven percent reduction in prefrontal gray matter) would be the result of environmental influence in the absence of some kind of trauma.

deal of conscious effort and that rarely produces the sort of fluency and pronunciation exhibited by native speakers. This is in part because adults who were not exposed to certain linguistic sounds (phonemes) during the critical period for language acquisition lose their ability to distinguish those alien phonemes from each other and from familiar ones (Werker and Tees, 1984).

A similar pattern appears in moral development. Particular cultures exploit a subset of the possible moral intuitions we are prepared to experience, much in the way that particular languages exploit a subset of the possible phonemes we are prepared to recognize and pronounce (Haidt, pg. 827). According to Shweder and his colleagues (1997), these intuitions cluster around what he calls the "big three" domains of human moral phenomena: the "ethics of autonomy" which concerns rights, freedom, and individual welfare; the "ethics of community" which concerns the obligations of the individual to the larger community in the form of loyalty, respectfulness, modesty, self-control, etc.; and the "ethics of divinity" which is concerned with the maintenance of moral purity in the face of moral pollution. Rozin and colleagues (1999) argue that these three domains correspond to three basic moral emotions: anger for autonomy, contempt for community, and disgust for divinity. Children develop different sets of intuitions/emotional responses depending on which moral "custom complexes" surround them. While all cultures appear to have practices associated with each of the big three moral domains to some extent, their emphases can be very different. Westerners, unlike the natives of Orissa, India, have relatively little familiarity with the ethics of divinity.

…In Orissa, India, many spaces and objects are structured by rules of purity and pollution. Foreigners and dogs may be allowed near the entrance of a temple complex, but only worshippers who have properly bathed may be allowed into the central courtyard (Mahapatra, 1981). In the inner sanctum, where the deity sits, only the Brahmin priest is permitted to enter. Private homes have a similar structure with zones of high purity (the kitchen and the puja room, where the household deity is kept), and lower purity (an entrance room, where visitors are entertained). The human body is given a similar structure, in which the head is the zone of highest purity, while the feet are highly polluting.

Children in Orissa constantly encounter spaces and bodies structured by purity, and they learn to respect the dividing lines. They learn when to remove their shoes, and how to use their heads and feet in a symbolic language of deference (as when one touches one's head to the feet of a highly respected person)… By participating in these interlinked custom complexes regarding the use of space and the purification of the body, children learn that a central project of moral life is the regulation of one's own bodily states as one navigates the complex topography of purity and pollution…. When such children later encounter the cognitive content of the ethics of divinity (e.g. ideals of sacredness, asceticism, and transcendence), their minds and bodies are already prepared to accept these ideas, and their truth feels self-evident… When an

American…travels in India, he may know how rules of purity and pollution govern the use of space, but he knows these things only in a shallow, factual, consciously accessible way; he does not know these things in the deep cognitive/affective/motoric way that a properly enculturated local knows them….

Social skills and judgmental processes that are learned gradually and implicitly then operate unconsciously, projecting their results into consciousness, where they are experienced as intuitions arising from nowhere (Haidt, 2001, 827-8).

Children in Orissa, India learn to see the world in terms of purity and pollution. Western children who are not exposed to the relevant culture complexes fail to develop such intuitions, but learn to develop intuitions of their own, some of which will seem alien to members of certain other cultures.

This learning process is largely emotional.

Perhaps the most powerful channel through which adults socialize children is by displaying their own emotional reactions. Shweder, Mahapatra, and Miller (1987) proposed a "social communication theory" in which children seek to understand the moral order that surrounds them and the "local guardians of the moral order" assist them in understanding it, largely by making affectively laden moral interpretations during the course of routine family and community life. Shweder et al. offer the

example of menstrual pollution among traditional Brahmins in Orissa, India, where a menstruating woman is expected to avoid physical contact with her husband, the kitchen, all sacred objects, and even her own children. When her child approaches, a menstruating woman might say, with some emotion, "I am polluted! Don't touch me! Don't touch me!" If menstruating aunts and sisters show similar reactions, and if temple priests react with distress or anger when a sacred site is defiled by contact with a person in a state of impurity, the child learns that there is a state of "pollution" in which touch becomes dangerous (Haidt, 2000).

To Western ears, talk of a menstruating woman's being "polluted" sounds primitive and silly, but to one who has developed intuitive emotional responses to such talk and the states with which it is concerned, it sounds like common sense. Likewise, members of other cultures may find our intuitions concerning things such as a woman's rights with respect to her husband utterly mysterious. What to us is a manifestly abhorrent and abusive relationship will strike normal, upstanding members of certain other cultures as perfectly acceptable.

As in the case of language, there appear to be critical periods for moral education. As noted above, patients who sustain ventromedial damage at a young age appear more likely to exhibit anti-social behavior than those who sustain such damage later in life (Anderson *et al.*, 1999). Rozin, Fallon, and Augustoni-Ziskind (1985) point out that children in cultures that do not emphasize matters of purity and pollution often develop such intuitions spontaneously

around the ages of seven or eight. (Think of children's frequent obsessions with "cooties.") These intuitions and concerns, however, tend to wither without cultural support. Something similar appears to be the case with respect to the ethics of autonomy and community. Around the age of four children tend to go from being relatively uninterested in matters of fairness to being obsessed with them, often overgeneralizing norms of fairness to inappropriate situations (Fiske, 1991). Minoura (1992) found striking differences in the socialization processes undergone by Japanese children of different ages living in America where their fathers were temporarily transferred for work. Of these children, those who spent a few years in America during the ages of nine through fifteen tended to develop American ways of interacting with friends and of reacting to and addressing interpersonal problems. The ones who spent time in America before the age of nine showed no such lasting effects, and the ones who arrived in America after the age of fifteen did not adjust as well to American life. Such late arrivals typically felt awkward behaving in American ways even while having excellent explicit knowledge of American behavioral norms.

Haidt (2001, pg. 828) summarizes the social intuitionist account of moral development as follows.

Moral development is primarily a matter of cultural shaping of intuitions. People can acquire explicit propositional knowledge about right and wrong in adulthood, but it is primarily through participation in custom-complexes (Shweder, *et al*., 1998) involving sensory, motor, and other forms of

implicit knowledge (Fiske, 1999; Lieberman, 2000; Shore, 1996), that are shared with one's peers during the sensitive period of late childhood and adolescence (Harris, 1995; Huttenlocher, 1994; Minoura, 1992) that one comes to feel, physically and emotionally (Damasio, 1994; Lakoff and Johnson, 1999), the self-evident truth of moral propositions.

3.6 The Illusion of Rationalist Psychology

In Sections 3.2-3.4 I developed an explanation for why moral realism appears to be true, an explanation featuring the Humean notion of projectivism according to which we intuitively see various things in the world as possessing moral properties that they do not actually have. This explains why we tend to be realists, but it doesn't explain, and to some extent is at odds with, the following curious fact. The social intuitionist model is counterintuitive.[37] People tend to believe that moral judgments are produced by reasoning even though this is not the case. Why do people make this mistake?

Consider, once again, the case of Mark and Julie, the siblings who decided to have sex. Many subjects, when asked to explain why Mark and Julie's behavior is wrong, engaged in "moral dumbfounding," bumbling efforts to supply reasons for their intuitive judgments. This need not have been so. It might have turned out that all the subjects said things like this right off the bat:

---

[37] That is, at least for Westerners.

"Why do I say it's wrong? Because it's clearly just wrong. Isn't that plain to see? It's as if you're putting a lemon in front of me and asking me why I say it's yellow. What more is there to say?" Perhaps some subjects did respond like this, but most did not. Instead, subjects typically felt the need to portray their responses as products of reasoning, even though they generally discovered (often with some embarrassment) that they could not easily supply adequate reasons for their judgments.

On many occasions I've asked people to explain why they say that it's okay to turn the trolley onto the other tracks but not okay to push someone in front of the trolley. Rarely do they begin by saying, "I don't know why. I just have an intuition that tells me that it is." Rather, they tend to start by spinning the sorts of theories that ethicists have devised, theories that are nevertheless notoriously difficult to defend. In my experience, it is only after a bit of moral dumbfounding that people are willing to confess that their judgments were made intuitively. Why do people insist on giving reasons in support of judgments that were made with great confidence in the absence of reasons?

I suspect it has something to do with the custom complexes in which we Westerners have been immersed since childhood. We live in a reason-giving culture.[38] Western individuals are expected to choose their own way, and to do so for good reason. American children, for example, learn about the rational design of their public institutions; the all important "checks and balances"

_____

[38] That is not to say that other cultures are not sufficiently "reason-giving" to produce this effect. I suspect, however, that Western culture is especially so.

between the branches of government, the judicial system according to which

accused individuals have a right to a trial during which they can, if they wish,

plead their cases in a rational way, inevitably with the help of a legal expert

whose job it is to make persuasive legal arguments, etc. Westerners learn about

doctors who make diagnoses and scientists who, by means of experimentation,

unlock nature's secrets. Reasoning isn't the only game in town, of course. The

American Declaration of Independence famously declares "these truths to be

self-evident," but American children are nevertheless given numerous reasons

for the decisions of their nation's founding fathers, for example, the evils of

absolute monarchy and the injustice of "taxation without representation." When

Western countries win wars they draft peace treaties explaining why they, and

not their vanquished foes, were in the right and set up special courts to try their

enemies in a way that makes it clear to all that they punish only with good

reason. Those seeking public office make speeches explaining why they should

be elected, sometimes as parts of organized debates. Some people are better at

reasoning than others, but everyone knows that the best people are the ones

who, when asked, can explain why they said what they said and did what they

did.

With this in mind, we can imagine what might go on when a Westerner

makes a typical moral judgment and is then asked to explain why he said what

he said or how he arrived at that conclusion. The question is posed, and he

responds intuitively. As suggested above, such intuitive responses tend to

present themselves as perceptual. The subject is perhaps aware of his "gut

reaction," but he doesn't take himself to have *merely* had a gut reaction. Rather, he takes himself to have detected a moral property out in the world, say, the inherent wrongness in Mark and Julie's incestuous behavior or in shoving someone in front of a moving train. The subject is then asked to explain how he arrived at his judgment. He could say, "I don't know. I answered intuitively," and this answer would be the most accurate answer for nearly everyone. But this is not the answer he gives because he knows after a lifetime of living in Western culture that "I don't know how I reached that conclusion. I just did. But I'm sure it's right," doesn't sound like a very good answer. So, instead, he asks himself, "What would be a good reason for reaching this conclusion?" And then, drawing on his rich experience with reason-giving and -receiving, he says something that sounds plausible both as a causal explanation of and justification for his judgment: "It's wrong because their children could turn out to have all kinds of diseases," or, "Well, in the first case the other guy is, like, already involved, but in the case where you go ahead and push the guy he's just there minding his own business." People's confidence that their judgments are objectively correct combined with the pressure to give a "good answer" leads people to produce these sorts of post-hoc explanations/justifications. Such explanations need not be the results of deliberate attempts at deception. The individuals who offer them may themselves believe that the reasons they've given after the fact were really their reasons all along, what they "really had in mind" in giving those quick responses.

Supposing that people engage in this kind of post-hoc fabrication may sound cynical, even paranoid, but there are good reasons to think that such phenomena are fairly common and not limited to people's explanations of their moral judgments. Maier found that subjects were much more likely to solve the problem he'd put to them (the solution involving tying together two ropes suspended far apart from the ceiling) if he'd casually swung one of the ropes while talking to them, but the subjects, when asked how they managed to solve the problem, were much more likely to attribute their insight to a salient but demonstrably useless cue, namely twirling a weight on a cord (Maier, 1931). Nisbett and Wilson (1977) argue that the subjects' introspective causal explanations given in this case and similar cases are often post-hoc fabrications .

> When asked to explain their behaviors, people engage in an effortful search that may feel like a kind of introspection. But what people are searching for is not a memory of the *actual* cognitive processes that caused their behaviors, for these processes are not accessible to consciousness. Rather people are searching for *plausible theories* about why anyone in their situation would have done what they did (Haidt, 2001, pg. 822).

Similar effects are seen in subjects whose actions were caused by means of post-hypnotic suggestion (Zimbardo, LaBerge, and Butler, 1993) and subliminal presentation (Kunst-Wilson and Zajonc, 1980). Such subjects readily devised

plausible but false explanations for their behavior. Perhaps the most dramatic illustration of this sort of effect is in the behavioral explanations given by split-brain patients, patients whose two cerebral hemispheres have been surgically separated in order to prevent the inter-hemispheric spread of dangerous seizures. Such patients will sometimes offer fabricated verbal explanations (produced by the left side of the brain) for actions performed by body parts under right-hemispheric control (Gazzaniga, 1985).

Whether or not this tendency toward believing in rationalist moral psychology is peculiar to Western culture or more universal is unclear. One place where we might expect this tendency to run rampant is among *philosophers*. As Robert Nozick (1993) wrote, "Philosophy means 'love of wisdom,' but what philosophers really love is reasoning." Deanna Kuhn (1991) has found that philosophers tend to engage in genuine reasoning where others, even others with domain-specific knowledge, do not. Thus, if all of this stuff about reasoning as a post-hoc fabrication sounds implausible to you, that may be because you belong to a rare class of people who really do arrive at moral judgments by reasoning. But my guess is that even among philosophers particular moral judgments are made first and reasoned out later. In my experience, philosophers are often well aware of the fact that their moral judgments are the results of intuition. As noted above, it's commonplace among ethicists to think of their moral theories as attempts to organize pre-existing moral intuitions.   The mistake philosophers tend to make is in accepting rationalism proper, the view that our moral intuitions (assumed to be roughly correct) must

be ultimately justified by some sort of rational theory that we've yet to discover. For example, philosophers are as likely as anyone to think that there must be "some good reason" for why it's okay to turn the trolley onto the other set of tracks but not okay to push the person in front of the trolley, where a "good reason," or course, is a piece of moral theory with justificatory force and not a piece of psychological description concerning patterns in people's emotional responses.

This tendency toward rationalist psychology has some important consequences which we will explore in the next chapter. Here's Haidt (2001, pg. 823) with a summary and preview:

> If moral judgment is generally a post-hoc fabrication intended to justify automatic moral intuitions, then our moral life is plagued by two illusions. The first illusion can be called the "wag-the-dog" illusion: We believe that our own moral judgment (the dog) is driven by our own moral reasoning (the tail). The second illusion can be called the "wag-the-other-dog's-tail" illusion: In a moral argument, we expect the successful rebuttal of our opponents' arguments to change our opponents' minds. Such a belief is analogous to believing that forcing a dog's tail to wag by moving it with your hand should make the dog happy.[39]

---

[39] There is a potential tension between my claim in this section that people tend to accept rationalist moral psychology and my claim in the last chapter (Section 2.4.2) that Smith's analytic naturalism attributes to people an unduly rationalistic

3.7 Conclusion

Moral judgment is, for the most part, driven not by moral reasoning, but by moral intuitions of an emotional nature. Our capacity for moral judgment is a complex evolutionary adaptation to an intensely social life. We are, in fact, so well-

conception of morality. I have two responses. First, here I claim only that certain cultures exhibit this affinity for rationalist moral psychology. Insofar as less rationalistic cultures employ moral concepts such as wrongness, the complaint about building rationalism into the meaning of moral terms is still legitimate. Second, even if all cultures were like Western culture, the point would still stand. Once again, we must distinguish between believing that someone can or has reasoned to some moral conclusion and believing that there is nothing more to the fact that some conclusion is true than that some group of people would, if fully rational and informed, arrive at that conclusion. People believe that moral judgments, like scientific judgments, are typically produced by episodes of reasoning. Hence people tend toward a belief in rationalist moral psychology. But at the same time, conceptually competent moralists need not believe that moral facts are merely facts about what conclusions we would reach if we more informed and more rational. People can accept a rationalist moral psychology without accepting a rationalist moral philosophy.

adapted to making moral judgments that our making them is, from our point of view, rather easy, a part of "common sense." And like many of our common sense abilities, our ability to make moral judgments feels to us like a perceptual ability, an ability, in this case, to discern immediately and reliably mind-independent moral facts. As a result, we are naturally inclined toward a mistaken belief in moral realism. The psychological tendencies that encourage this false belief serve an important biological purpose, and that explains why we should find moral realism so attractive even though it is false. Moral realism is, once again, a mistake we were born to make.

It is our biological inheritance to see the world in moral terms, to "gild and stain" the world with moral colors. But each culture paints its own design and with its own palate. An individual's moral sensibility is shaped by the "culture complexes" that surround her. Morality, like language, is universal in its form, but local in its content.

Moral realism is not the only illusion that plagues our moral thought. We Westerners, and perhaps others as well, are inclined toward a mistaken rationalist moral psychology, the belief that our moral judgments are caused by moral reasoning. We fail to appreciate the intuitive nature of our moral judgments, mistakenly believing that the rational tail wags the emotional dog. As a result we misunderstand not only ourselves (the "wag-the-dog" illusion) but others as well (the "wag-the-other-dog's-tail" illusion).

Amidst all of this misunderstanding of the moral mind and the moral world, we should hardly be surprised if the way we think and talk about moral matters is

less than ideal. Indeed, I believe we can do better. Part II of this essay is

devoted to explaining how this might be accomplished.

Part II:

What to Do About It

Chapter 4

Practical Anti-Realism

In Part I we concluded that moral realism is false and that moral psychology provides us with a good account of our mistaken inclinations toward realism. Having embraced anti-realism from a *theoretical* standpoint, we must now consider what it means to embrace it from a *practical* standpoint. Given that there is no fact of the matter about moral right and wrong, *what are we to do*? Is morality just a bunch of nonsense? Now that we know the terrible, horrible truth about morality, is there any sense left in trying to lead a moral life?

Of course, there is. Morality$_2$ survives in the absence of morality$_1$, which is to say that we are perfectly free to go on caring about each others' well being and guiding our actions accordingly, knowing all the while that our actions are neither right nor wrong. But, at the same time, it would be foolish to go on just as before. The common sense approach to moral life, the way we ordinarily think and talk about moral matters, has been shaped by psychological tendencies and cultural practices that are thoroughly realist. While there is certainly nothing *wrong$_1$* with doing things the same old way, the conclusions we've reached suggest that at least some aspects of our established moral practices are likely to be counterproductive. In this chapter we begin the process of separating baby from bath water.

4.1 Some Preliminaries

4.1.1 Error Theories And Pessimism

One might view the enterprise of this essay as a salvage operation, an attempt to

face squarely the demise of moral realism and make the most of what remains.

In a sense that's correct, but the tone is all wrong. Error theories and anti-

realisms are typically taken to be *pessimistic* positions because they tell us that

we were *wrong*, *naïve*, that the world is not as we believed or hoped it to be.

They are viewed as failures, theories of last resort (McNaughton, 1998, pg. 98). I

take a different view. If the goal of philosophy is to discover and resolve esoteric

philosophical problems ("Do numbers exist?"), then finding out that you and your

common sense were right all along is an unmitigated success ("Yes! They do!").

If, instead, the goal is to gain insight into problems that are primarily the concern

of people outside of philosophy (war, famine, social injustice, that sort of thing),

then finding out that you and your common sense were right all along is a cause

for *disappointment*. It means that we've discovered nothing beyond that which

ordinary people take for granted. (Headline: "External World Exists!

Epistemological Skepticism Defeated!") An error theory, on the other hand, is a

cause for *optimism*. In a world full of practical problems, news of past mistakes is *good news*, the first step toward avoiding them in the future.[1]

---

[1] Of course there is value in epistemological conservatism and respect for common sense, but I think that philosophers often overvalue these things. New ideas are more likely to be useful than old ideas simply because old ideas have already been put to use or discarded. At the same time, though, new ideas are more likely than old ideas to be incorrect. (If it's such a good idea, why didn't someone think of it before?) Thus, there is a tension between the theoretical and practical standpoints with respect to the evaluation of new ideas. New ideas are more likely to be wrong, but at the same time they're our best hope for progress. What to do? The following case illustrates what I take to be a good balance. Suppose there is a large body of evidence suggesting that a certain disease, one that has proved very difficult to cure, is caused by a virus. A new study suggests that this disease is in fact caused by environmental pollutants. The proper response to the new study would be a mixture of skepticism and hope. In the face of much conflicting evidence, we will have our doubts about this new study and do our best to explain the data it reports in a way that is consistent with the viral theory. But in our hearts we will hope that the conclusion supported by the new study is correct, that we had been wrong all along, and that we are now in a position to make some progress. I believe that a similar mixture of skepticism and hope is an appropriate response to the revisionary meta-ethics outlined in Part I.

4.1.2 The Baggage of Moral Realism

Some error theories are more unsettling than others. If it turns out that common sense is wrong about, say, the nature of causation, that conclusion, however unwelcome it may be to epistemological conservatives, is not likely to upset the folks back home. Meta-ethical error theories of the kind advocated by myself and the late John Mackie, however, are more likely to disturb, say, a well-meaning relative who innocently inquires after the subject of my dissertation: "No such thing as right and wrong? This is what they teach you? I suppose you think we should let Osama Bin Laden do as he pleases!"

Such reactions reveal a misunderstanding encouraged by the dominance of moral realism. Under a typical realist regime, it's wrong to oppose something

---

Another worry: Does this mean that any usable news is good news? Should we say that the news that you have cancer is *good news* since now, having heard the news, you're in a position to start treating it? What counts as good or bad news depends on your frame of reference. To learn of your cancer is to discover that you have a problem that you didn't know you had. In contrast, the problems that I claim are exacerbated by our belief in moral realism are problems that we *already know we have*. Thus, finding out that we have relevant false beliefs introduces no new practical problems, only possible solutions, which is why I say that the rejection of moral realism is good news.

that is not wrong. ("But I haven't done anything wrong!" cries he against whom undue punitive actions have been taken.) Combine my claim that no action is wrong, with the realist's background assumption that it's wrong to oppose actions that are not wrong, and we have a recipe for disaster (as well as a contradiction): If terrorists want to kill thousands of innocents, who are we to stop them? After all, they're not doing anything *wrong*!

In rejecting moral realism we must reject the whole thing, including the assumption that it's wrong to oppose that which is not wrong. This should go without saying since, as noted above, it can't be the case that no actions are wrong and that certain acts of opposition are wrong. Nevertheless, "X is not wrong," and "Allow X!" are so closely allied in ordinary ethical discourse that we are likely to assume a parallel alliance between the meta-ethical claim "Nothing is wrong," and the libertine dictum, "Allow everything!" Conflating anti-realism or error theory with universal permissiveness or nihilism (See Section 1.5.4) is a mistake to be avoided as we consider the practical options available to anti-realists and error theorists.

Avoiding this pitfall is easy once it has been pointed out, but there is other realist baggage with which it is more difficult to part. We return once again to the distinction introduced in Chapter 1 between:

moral₁: of or relating to the facts concerning right and wrong, etc.

and

moral$_2$: of or relating to serving (or refraining from undermining) the interests of others.[2]

One of the central claims of this essay is that we can retain the benefits of morality$_2$ while rejecting morality$_1$. Obviously this would not be possible if morality$_1$ and morality$_2$ were inseparable, and therefore I will take a few moments to defend this all-important distinction.

In Chapter 1, I argued for the plausibility of this distinction using the example of Brett, an individual who is highly moral$_2$ (showing great concern for the interests of others), but not at all moral$_1$ (explicitly denying that there is any fact of the matter concerning what one ought to do). I described Brett in a bit more detail and asserted that there is no incoherence in Brett's psychology as described, and therefore no problem with disentangling the moral$_1$ from the moral$_2$. One obstacle to drawing this distinction is the possibility that moral realism might be true. If there were a fact of the matter about what's right and wrong, then it would be possible for Brett to be inadvertently moral$_1$ in spite of his avowals to the contrary. However, we've since concluded that realism is false and thus eliminated this particular challenge to the moral$_1$/moral$_2$ distinction.

---

[2] In light of our discussion of the "ethic of divinity" this definition is actually overly restrictive, reflecting my Western bias. However, for the purposes of this essay this should not matter.

There is, however, a different objection we should consider. One might think that it is, as matter of psychological fact, simply impossible to behave morally$_2$ if one does not, on some level, accept moral realism. The thought may be that human beings simply will not be motivated to act morally$_2$ if they do not believe that there is a fact of the matter concerning how they ought to behave. This issue concerning the relationship between motivation and moral judgment is essentially the converse of the more familiar issue surrounding the thesis of *internalism* (Smith, 1994). Internalism is the view according to which moral judgments are necessarily motivating (at least insofar as agents are rational). According to internalism a rational agent cannot, for example, sincerely judge that some action is morally wrong without being motivated to refrain from performing it. Internalism, though doubtful in my opinion, is not completely implausible. After all, someone who performs an action that she professes to be wrong invites us to doubt her sincerity (or, alternatively, the strength of her will). However, the thesis behind the present objection to the moral$_1$/moral$_2$ distinction is far more radical than internalism. According to this view, it's impossible to exhibit stable moral$_2$ behavior and patterns of thought without believing that failure to do so would or could be immoral$_1$. This is akin to the belief, common within certain theistic subcultures, that atheists, people who are literally not God-fearing, cannot be trusted to behave morally$_2$. Few of us would believe *that*, so why would one think that moral$_2$ behavior requires some other kind of metaphysical belief?

There are, I think, two paths to this conclusion. The first is by way of thinking that there is little more, perhaps nothing more, to being a moral realist than living a moral$_2$ life. This path is more likely to be taken by contemporary realist philosophers, and parallels the rather eccentric belief that anyone who exhibits morally worthy behavior is a *de facto* theist. The second path is by way of a thinking that full-blown realist convictions are really very important for living a moral$_2$ life. This parallels the thought of theists who do not trust atheists.

Following the first path, someone might think that being a moral realist involves nothing more than making moral judgments through a process of rational reflection.[3] When, for example, Brett considers the interests of others in deciding how to act, he is essentially asking himself, "What sort of behavior can I endorse as one rational agent among many?" The decisions that emerge from such reflections, at least when they are successful, are judgments that not only reflect the agent's personal interests, but also a more general interest in acting in ways that a rational agent, perhaps all rational agents, can endorse. And insofar as that's the case, what agents such as Brett are really doing, in spite of their avowals to the contrary, is aiming to do *the right thing*. They are moral realists in spite of themselves.

While some may find this sort of account compelling, I simply fail to see why engaging in rational reflection of a practical nature amounts to moral realism. It's one thing to employ one's capacity to reason in deciding what to do, another to think that the practical question one faces has a *correct answer*—that is, an

---

[3] There are strains of such a view in Korsgaard (1996).

answer that is correct for reasons that do not depend on the particular desires or interests one happens to have. Of course, it depends on what one means by "rational." If, on the one hand, being rational is more or less a matter of being logically consistent, a matter of being "thinly rational," then it's just false to say that one cannot make rational decisions that show concern for the interests of others without believing that there is a fact of the matter concerning whether those decisions are morally correct. What about more substantive notions of rationality, so-called "thick rationality?" Such notions are irrelevant here. Holding out for some more substantive notion of rationality might save internalism, but it's beside the point in this case. Brett, and people in general, need not be thickly rational. My claim is only that it's possible to be morally$_2$ motivated without being a moral realist. If it turns out that one can't be motivated by concerns for the interests of others *and exhibit a certain kind of substantive rationality* without being a moral realist, then so be it. This is just an another instance of a familiar problem with rationalist ethics. The more substantive morality one packs into "rationality" the more optional "rationality" becomes.[4]

The first path to conflating morality$_1$ and morality$_2$ makes the issue a conceptual one. From the armchair one (fallaciously) deduces that being moral$_2$ is just a matter of being rational. The second path, in contrast, depends on a bold empirical claim about moral psychology, and now, having considered some relevant moral psychology in Chapter 3, we are in a position to handle this objection. Once again, we are considering the claim that one cannot be reliably

---

[4] See Section 2.4.2.

motivated to behave morally$_2$ unless one is a moral realist, one who believes in morality$_1$. Our discussion in Chapter 3 suggests that morality$_2$ precedes morality$_1$ in almost every way. The social impulses that give rise to morality$_2$ were in play long before there were people and long before anyone ever bothered to ask, "But is there really any fact of the matter about what's right or what's wrong?" Chimpanzees certainly exhibit moral$_2$ behavior,[5] but it would be a bit of a stretch to say that chimpanzees are generally moral realists. Likewise, it would be a stretch to call them anti-realists. The right thing to say about them is that they are neither realists nor anti-realists. For them, the meta-ethical question simply never arises. But if that's right, then, at the very least, it's possible for chimpanzees to be morally$_2$ motivated without being moral realists. Are we humans unable to do the same?

Perhaps one will say that insofar as one is capable of raising the meta-ethical question, one cannot be morally$_2$ motivated without being a moral realist. But why should one say that? A creature develops the cognitive capacity required for raising certain kinds of philosophical questions, and then, all of a sudden, many of his most basic social impulses, developed and refined over thousands, even millions, of years of evolution, are thrown in jeopardy as he wavers between philosophical conclusions? Absurd! Chimpanzee mothers care for their young not because they believe that doing so is morally required of

---

[5] That is, behavior "of or relating to serving (or refraining from undermining) the interests of others." Such behavior may not be "moral" in the ordinary sense since the ordinary sense conflates morality$_1$ and morality$_2$. See Section 1.2.

them, but because that's what chimpanzee mothers do. A human mother can ask, "Do I really have a moral obligation to care for my child?" but the fact that she can ask this question doesn't mean that her basic motivation for raising her children is radically different from those of a chimpanzee. A human mother who asks this question in a spirit of rational inquiry[6] is not at all likely to abandon her children if she ultimately comes to the conclusion that she has no such obligation. Fortunately for us and our genes, the emotional dog wags the rational tail.

A nice illustration of this realist-rationalist tendency to turn philosophical inclinations into universal principles of moral psychology is the discussion among moral psychologists surrounding the "Raskolnikov regression" (Fishkin, 1984).[7] Lawrence Kohlberg (1969, 1984), the archetypal rationalist moral psychologist and moral realist, is known for his theory of moral development according to which individuals pass through an invariant sequence of stages of moral development.   One of the earlier stages of moral development is Stage 2, characterized by a "naively egoistic orientation" (1984, pp. 43-4). A later stage, not attained by all, is Stage 5, characterized by a "contractual legalistic orientation." (The final stage, attained by only a select few, is Stage 6 in which one bases one's moral judgments on "principles of choice involving appeal to logical university and consistency" (1984, pg. 44).) Kohlberg was surprised to

---

[6] As opposed to, say, while in search of a justification for acting upon her preexisting frustrations with motherhood.

[7] My thanks to Joshua Knobe for bringing this issue to my attention..

find that many college students who were previously identified as Stage 5 moral reasoners were suddenly espousing a "relativistic" or "subjectivist" moral outlook which, according to Kohlberg's scheme, places them back in Stage 2, a stage typically associated with young children. Kohlberg struggled in a number of ways to account for this phenomenon in which individuals who once exhibited fairly sophisticated (i.e. realist and proto-rationalist) moral thought could regress back into putatively child-like moral reasoning (Fishkin, 1984, 162-8).   According to Fishkin (pg. 167), none of these moves is satisfactory.

> Subjectivism thus constitutes a basic challenge to the root assumptions of the Kohlberg theory. This challenge derives from the fact that Kohlberg's classifications are primarily normative ethical but the phenomenon of subjectivism is essentially meta-ethical. There is no place in Kohlberg's scheme for meta-ethical questioning except for the gaps between stages.

Kohlberg's claim that a "subjectivist" or "relativist" cannot be a moral adult are strikingly similar to those made by realists who would insist from their armchairs that one cannot behave like a moral adult, i.e. morally$_2$, in the absence of moral$_1$ commitments of the sort that define Kohlberg's later moral stages.

The human social instincts that undergird our commitment to morality$_2$ run deep. Our meta-ethical views concerning the truth of moral realism are not without consequences—or so I shall argue—but their effects are marginal compared to those that flow from our more basic social natures. Morality$_1$ is not

only distinct from morality$_2$, it is, according to the argument in Section 3.4, a by-product of our efficiently implemented capacity for moral$_2$ judgment.

4.1.3 Objective and Subjective Evaluation

I've denied that there is any fact of the matter concerning what we *ought* to do, and yet I am in the process of developing a practical proposal for how we should think and talk about moral matters. What, then, am I doing in making this proposal? Am I not saying that my proposal is *right*? *Better* than its alternatives? A proposal we *ought* to take seriously? Have I not denied the possibility of meaningful evaluation only to presuppose it?

　　No. And the reason should now be clear, though it bears some brief amplification. What I claim is that there is no fact of the matter about what one ought to do, which actions are right or wrong, which states of affairs are better than others, etc. This does not mean, however, that there are no facts concerning which actions and states of affairs are favored by our values and/or desires, facts which in turn may make for meaningful evaluation using words like "ought," "right," and "better." To take an instance of the simplest sort of case, if I want soup and not salad, then, *ceteris paribus*, I ought to order the soup; my ordering the soup is the right thing to do; and my receiving the soup is better than my receiving the salad. The "ought" here is "hypothetical" not "categorical," as are the "right" and "better." These terms, as I use them, indicate judgment

220

according to a subjective standard, one grounded in the desires and values of particular humans such as myself, and not an objective standard. I will argue later (in Section 5.2) that some of these evaluative terms are more problematic than others, and I will attempt to avoid the problematic ones so far as this is possible. For now, I simply wish to make it clear that, to the extent that I use such terms, they reflect imperatives that are merely "hypothetical" and evaluations that are ultimately subjective.

4.1.4 Analytic Naturalism Revisited

In Chapter 2 we considered several versions of analytic naturalism and rejected them on the grounds that each of them, in its own way, distorts the meanings of moral terms. Having discussed a bit of moral psychology, we are now in a position to give a more unified account of why these theories fail. At the heart of any version of analytic naturalism is an account of what certain moral terms mean, and what a term means is partially, if not entirely, a psychological matter. The accounts of moral language and concepts offered by analytic naturalists are inevitably esoteric—recall that they must be *indirect* to dodge Moore's Open Question Argument—making reference to idealized rational agents, variable moral frameworks, etc. We might consider, then, whether or not these esoteric accounts of moral language/concepts are compatible with the projectivist/social intuitionist account of moral psychology developed in the last chapter.

Let us begin with Michael Smith's view, the most explicitly rationalist form of analytic naturalism we've considered. According to Smith's account (slightly modified), what we mean when we say that an action is "wrong" is that it's something of which we would disapprove if we were fully informed and fully rational. In the last chapter (Section 3.4) I compared judgments of moral wrongness to judgments of sexiness, arguing that they are comparable instances of affectively laden projection. If your heart beats more than once a month you know that "sexy" does not mean "that with which I would desire to engage sexually if I were fully informed and fully rational," and, if the comparison between sexiness and wrongness is a fair one, then Smith's analysis is comparably implausible. Once again, when it comes to moral psychology, the emotional dog wags the rational tail, not the other way around. Rarely do we make moral judgments through the deliberate application of moral principles to particular circumstances. Rather, we simply see the world in moral terms, "gilding and staining" the world with "colours borrowed from internal sentiment." The thought that "Incest is wrong!" is much closer to[8] "Incest! Yuck!" than it is to "Incest: Something of which we would disapprove if we were fully informed and fully rational." To find something morally repugnant one need not have any thoughts

---

[8] "Closer to," but not "equivalent to." One can say without contradiction that one finds incest thoroughly disgusting but believes that it's not wrong. My point, rather, is that the thought that incest is wrong *tends* to be grounded in an affective response, one involving a certain kind of morally-laden disgust.

(explicit or implicit) about fully informed and fully rational individuals and their proclivities.

To drive this point home, consider the chimpanzees who displayed their moral (proto-moral?) indignation in punishing a pair of adolescent group members for having delayed the group's feeding the evening before.[9] It's hard to believe that those angry chimps were thinking anything even close to, "Those two adolescents have done something of which we would disapprove if we were fully informed and fully rational!" (Chimps may be smart, but they don't read Kant.) Therefore, the only way that Smith's account of moral concepts could be correct for humans is if human moral judgment is radically different from chimpanzee moral judgment, a matter of drawing conclusions about the desires of idealized rational agents rather than responding emotionally and intuitively to certain kinds of behavior. This is very doubtful. The evidence strongly suggests that human morality is emotional and intuitive, with its roots in primate morality, and for this reason we should be skeptical of any theory of moral meaning that turns basic moral thought into an esoteric affair.

The same objection applies to the other forms of analytic naturalism we have considered. It's similarly unlikely that "wrong" means "having the property that plays the wrongness role according to mature folk morality"[10] in the minds of either chimps or humans. Given the intuitive and emotional nature of moral judgment, one should be able to think of something as morally wrong without

---

[9] See Section 3.5.

[10] Recall the discussion of Jackson's view in Section 2.4.3.

having any thoughts about which properties play which roles according to idealized versions of our current moral theories. Likewise, the psychology of moral judgment suggests that one can have thoughts about right and wrong without having such esoteric thoughts as those concerning that which we would desire to desire under conditions of "full imaginative acquaintance."[11]

What we know about moral psychology makes similar trouble for relativist analytic naturalism, according to which "wrong" means, roughly, "wrong according to our moral framework" (Harman, 1996). In many cases at least, when one sees something as wrong, one projects wrongness onto it and thus sees that something as having a mind-independent property of wrongness. One does not see the wrongness as coming from oneself or as flowing from a particular society's norms any more than one sees sexy things as sexy in virtue of one's psychological projections or the content of local standards of sexiness. When one is struck by the apparent wrongness of an act of, say, incest, one's thought is not first and foremost, "That is disallowed by the norms that happen to be in place here and now!" For most people, the thought is, once again, much closer to "Yuck! That's just plain *wrong*!" where the "just plain" suggests, among other things, that the observed wrongness is absolute and not relative. The wrongness of incest need not strike the conceptually competent as any more culturally relative than the greenness of leaves. In light of the projective nature of moral "perception," it would be strange if the sort of sophisticated detachment that relativism requires were built into the concepts that give rise to basic moral

---

[11] See discussion of Lewis' view, Section 2.4.3.

thought in humans, and this would be stranger still in the case of our nearest living relatives.

4.2 Assessing Our Moral Practices

What do we want from morality? If we are moral realists, part of what we want is knowledge of Moral Truth, both for its own sake and as a means to living rightly. Having rejected moral realism, we must ask what *else* we might want from morality. One answer, the radical nihilist's answer, is "Nothing." But radical nihilism is not a live option for most of us because most of us, even upon rejecting moral realism, still want to live in a moral$_2$ world, a world in which people care about one another's well-being and regulate their behavior accordingly. Maintaining such a world requires the existence of widely-observed norms for behavior and an ongoing discussion concerning how those norms should evolve in the future. To a realist, these norms are, at least in some cases, approximations of the *correct* norms. To an anti-realist, they are simply norms, guidelines for behavior that perform a function.

What function? Surely there is no single answer. Some norms perform functions that are not generally useful, e.g. keeping power and privilege in the hands of a few. But this is not the case for most moral norms, or so we assume. Regardless of what moral norms actually do, we[12] would like them to promote

_____

[12] Most of us, anyway.

happiness, freedom, respect for others, honesty, peace, understanding, etc. How exactly these ends are to be furthered and what tradeoffs among them are desirable and/or tolerable is a matter of much dispute, perhaps *the* matter of much dispute. Nevertheless, we can say that the promotion of these ends in general is the practical function of morality and that, in the absence of moral realism, the value of our moral practices is to be assessed solely in terms of how well they perform this function.

Our moral language is thoroughly realist. We speak as if moral facts are out there waiting to be discovered: "Does a fetus have a right to life?" "Is it wrong to kill animals for food?" "What obligations do we have toward the poor?" This comes as no surprise given that most people are moral realists, which itself is no surprise given the projective nature of moral psychology. But does it make sense to go on talking and thinking like moral realists if there really is no fact of the matter about what's right and wrong? The surprising answer given by most philosophers who have considered the issue is *Yes*. Once again, Geoffrey Sayre-McCord (1988, pg. 2) tells us that, "The plausibility of anti-realism… depends largely on preserving our normal ways of speaking even while challenging the natural (though perhaps naïve) realistic interpretation of what is being said." Sayre-McCord not only thinks that it makes sense to talk like a realist after rejecting realism. He thinks that our rejection of realism should be *contingent* on our being able to go on talking like realists! Simon Blackburn (1984, 1993a) suggests that we become *quasi-realists*, "quasi-" because we know that moral properties are mere projections and "realists" because we

nevertheless think and talk just like realists except, perhaps, in moments of meta-ethical reflection. Richard Joyce (2001) suggests that we become *fictionalists*, recognizing that moral realism is false while maintaining the fiction that it is true through our language and, to a large extent, our thought. Even John Mackie (1977) appears inclined toward something like fictionalism, speaking like a realist when it suits him (Blackburn, 1993a, pp. 149-50). Some anti-realist philosophers describe themselves as *constructivists* (Rawls, 1971, 1980). They deny that there are full-blown moral facts, but argue that there are moral "facts" that may be "constructed" according to a certain rational procedure that may serve as full-blown facts for practical purposes. Thus, according to this view as well, there is nothing wrong with our realist moral talk.

A wholly different anti-realist option, mentioned above, is *nihilism* (or, to be safe, "radical nihilism") by which we would give up on morality altogether. While this is not a serious option for those of us who appreciate the extremely important function that morality (i.e. morality$_2$) performs, it's an important foil in the present discussion. I've suggested that realists fail to take anti-realism seriously because they think it leads to nihilism, and anti-realists appear eager to preserve our moral practices in their present form for fear that in abandoning them we will have no option other than nihilism.

I believe that the best option, the one most likely to further our ultimate aims as moral$_2$ people, is none of these. Rather than faking it (fictionalism, quasi-realism), remaking it (constructivism), or forsaking it (nihilism), we would be wise to modify our moral practices so that they give us the fullest possible moral$_2$

benefit while avoiding the ignorance, confusion, stubbornness, and unnecessary conflict that arise out of our adherence to moral practices that are grounded in false metaphysics. The key to understanding why our current moral practices are so fraught with problems and how these problems might be avoided lies in moral psychology.

4.3 Moral Psychology and the Scale Problem

First, let us briefly review the key points made in the last chapter.

Moral judgment is for the most part intuitive and emotional. We can offer reasoned justifications on behalf of our moral judgments, and typically do when our judgments are questioned, but the judgments themselves are, for the most part at least, not produced by episodes of reasoning, although people tend to think that they are.

The phenomenology of moral judgment is misleadingly realist. We see the world as having mind-independent moral properties, there to be perceived though our "moral sense," properties like the inherent wrongness of incest. Our experience of the moral world is as such because moral judgments are "easy." They are produced by automatic, unconscious, and intuitive processes akin to those that allow us to immediately and reliably sort human faces by sex. The automatic nature of these processes results in an experience of *seeing*, as opposed to *figuring out*, which things are right or wrong.

A human moral sensibility is both innate and inculturated. Like our capacity for language, it is an emergent phenomenon. The basic structure of human morality is innate, but the "custom complexes" in which one is immersed play a crucial role in shaping the content of one's moral sensibility. Thus, moral development is a natural process in which one's biological dispositions and surrounding culture interact to produce an individual who sees the world in a particular moral light.

Such is the nature of human morality, and one can see why individuals living in small communities with others who are similarly disposed would flourish. To begin, the moral system that emerges from this psychology is flexible enough to allow for advantageous cultural adaptation. For example, in conditions of relative abundance a culture might do well to discourage cannibalism under all circumstances, thus avoiding unnecessary opportunities for the spreading of disease and removing some nutritional incentives to violence. In conditions of relative scarcity, a culture might do well to allow some instances of cannibalism, thus permitting some individuals to survive when they otherwise might not. Of course, flexibility is not an unlimited good. One can imagine the social disaster that would result from each individual's having a propensity to fashion his own code of behavior to suit his personal circumstances. Fortunately, human moral psychology does not make for radical individualism. Individuals learn the local morality much as they learn the local language.

Thus, human moral psychology affords an excellent balance of social benefits to individuals living in small communities. There is enough flexibility at

the community level to allow communities to develop norms for behavior that suit their circumstances and enough flexibility within individuals to allow them to adapt to life in whatever sort of community they happen to be born into. At the same time, this individual flexibility and the community-level flexibility that emerges from it operate within limits. In any behaving system there is always a trade-off between flexibility and efficiency.[13] Once a developing human has had enough time to figure out how life is lived in her community, there is little to be gained from remaining flexible in her range of social behaviors and much to be gained from having those behaviors become automatic and intuitive. Once again, there is a natural analogy with language. After a Chinese child has had enough exposure to her native language to understand its basic structure, that child is better off devoting her cognitive resources to becoming an efficient speaker of Chinese rather than a possible of speaker of many languages she will never need to speak. When it comes to coordinated social behavior, either in the case of language or adherence to social norms, flexibility followed by efficiency is highly adaptive.

At least in small communities. What small communities provide, among other things, is relative homogeneity. Homogeneity makes for clear signals to children, allowing new members of the community to pick up the local culture

---

[13] The machines we make perform certain jobs better than we can by ourselves, but each machine is limited in the jobs it can perform. Humans, in contrast, can do an incredible number of things reasonably well, but increasingly few things as well as the machines we've devised to perform those tasks.

quickly. It also makes for smooth interactions among adults who have developed compatible behavioral dispositions as a result of having learned those behaviors from similar environments—different houses on the same street, so to speak. The adults, of course, are somewhat set in their ways, but that's not a problem because they're all set in roughly the same way and their day-to-day interactions do not involve people outside of their community.

But what about larger communities? As communities grow they become increasingly heterogeneous because of their sheer size and because of the diversity of roles that individuals and groups can play within larger communities. Opportunities for division of labor and specialization create a diversity of circumstances and subcultures, while technological advances in transportation and communication bring individuals from different cultures into contact with one another. Thus, the development of human civilization has, for better or for worse, disrupted the relative homogeneity of circumstances enjoyed by our ancestors.

But it has done so in an incomplete way. Above I noted two effects of homogeneity, the ease with which it allows children to become enculturated and the ease with which it allows enculturated adults to get along with one another. An essential feature of our current circumstances is that a modern child's environment, while not nearly as homogeneous as environments of the past, is vastly more homogeneous than the world with which that child will ultimately interact. Children aren't raised by multicultural committees. They're raised by one or two parents whose values are likely to be fairly similar to one another and in a community that is far from representative of their world at large. Parents

who make an effort to expose their children to different cultures perform a valuable service for their children and others, but doing this does not produce children that are "omnicultural." The limitations of the developing human mind make it inevitable that any given child will be a product of a particular culture or particular set of cultural influences.

In the past, people raised in a particular culture had relatively little to do with cultural outsiders, but that is far from the case today. Christian fundamentalists live with abortion laws that they would not dream of enacting themselves. Gay couples are denied the economic benefits of marriage because their lifestyle is unacceptable to the majority of their compatriots. People die in wars they do not support. Children in the developing world die of starvation because people in the developed world choose not to help them or create circumstances that make it impossible for them to help themselves. People in the developed world occasionally die when the frustrations of the developing world find expression in the form of terrorist attacks. Today, an individual's sphere of causal influence extends far beyond the environment that is primarily responsible for shaping her moral sensibility, and as a result ours is a world of people who see the world in radically different moral colors attempting to impose their respective wills on one another. In other words, the Nuclear Age has arrived, and we're still using Stone Age moral psychology. The human tendency to absorb the values of one's immediate environment and project those values onto the entire world may work well enough for life in small, relatively isolated hunter-gatherer bands, but it's absolutely disastrous for billions of people raised

in a variety of different cultures and subcultures who must share a world in spite of their incompatible worldviews. Human moral psychology doesn't scale well.[14]

4.4 Moral Realism and the Revisionist Proposal

Moral realism is the theoretical expression of the Stone Age moral psychology with which we are saddled. It is an illusion that exacerbates conflict and promotes misunderstanding. If I want things to be one way and you want things to be some other way, we might be able to reach some sort of compromise. But, if I want things to be one way, and if I believe that the way I want things is not *merely* the way I want things but also the way things *ought to be*, and if I believe further that it's *just plain to see* that my way is *the way things ought to be* and that anyone who says otherwise must be *outright lying* or *willfully refusing to see the truth*, and if you want things to be some other way and you're just as

---

[14] I do not wish to suggest any sort of "group selectionism" according to which human moral psychology evolved to make life go well for small groups of people. My claim is simply that insofar as human moral psychology does work well for small communities, it does not work nearly as well in the modern context.

I also do not wish to suggest that all scaling up leads to trouble. The case for the scale problem in morality rests on the details concerning moral psychology and not on some generic argument of the form "We no longer live in the environment in which X evolved; therefore we must change X."

convinced of the *rightness* of your position as I am of mine, then what chance do we have of reaching a reasonable compromise? I see you as an errant child, someone who has lost the way, someone who wasn't paying attention on the day right and wrong were explained, or, perhaps, as someone who was paying attention but who, for whatever inexplicable reason, has chosen to cast aside what is right and good in favor of that which is base and evil. And you, of course, see me in a similar light. I attempt to argue with you and am amazed at your obtuseness. The words I speak so clearly reveal the truth, and yet you persist in your wrongheaded ways. And you are similarly perplexed by me and my stubbornness. Haidt (2001, pg. 823) summarizes the social-intuitionist take on ordinary moral discourse:

> The bitterness, futility, and self-righteousness of most moral arguments can now be explicated. In a debate on abortion, politics, consensual incest, or on what my friend did to your friend, both sides believe that their positions are based on reasoning about the facts and issues involved (the wag-the-dog illusion). Both sides present what they take to be excellent arguments in support of their positions. Both sides expect the other side to be responsive to such reasons (the wag-the-other-dog's-tail illusion). When the other side fails to be affected by such good reasons, each side concludes that the other side must be insincere, closed-minded, or even devious…. In this way the culture wars over issues such as homosexuality and abortion can generate morally motivated players on both sides who

believe that their opponents are not morally motivated (Haidt & Hersh, 2001, Robinson, Keltner, Ward, & Ross, 1995).

A mess, indeed. But how to clean it up? Must we resign ourselves to a world of endless conflict and misunderstanding? Haidt (Pg. 823) suggests a shift in tactics:

> Moral reasoning may have little persuasive power in conflict situations, but the social intuitionist model says that moral reasoning can be effective in influencing people before a conflict arises… If one can get the other person to see the issue in a new way, perhaps by reframing a problem to trigger new intuitions, then one can influence others with one's words.

This, however, does not get to the heart of the problem. It is a gesture toward more subtle, less explosive forms of moral warfare, not peace. Moreover, it is a lesson that most professional moral communicators have already learned from experience. No surprise that novels, plays, metaphors, and anecdotes are more effective means of propaganda than philosophical arguments and statistics. (I've never stayed in a hotel room that came furnished with a copy of Kant's *Grundlegung*.) As Haidt suggests, a better understanding of moral psychology may be used to further one's own moral agenda—a good or bad thing depending on the agenda in question. But I propose instead that we use our understanding

of moral psychology to transcend our ordinary modes of moral discourse rather than to operate more effectively within them.[15]

Once again, the enemy, the wolf in sheep's clothing, is moral realism. Conflicts of interest may be inevitable, but they need not be exacerbated by people's unflagging confidence that they're *right* and that their opponents are *wrong*. The solution, then, is to get rid of realist thinking and to start by getting rid of realist language. Speak only in terms that make the subjective nature of value plain. Instead of saying that capital punishment is *wrong* say that you are opposed to it. Say that it is an ineffective deterrent, difficult to implement in a colorblind fashion, and likely to lead to irreversible mistakes. And then *say no more*. Instead of saying that eating animals is *wrong* and a form of *murder*, say that you are opposed to eating animals because you wish to alleviate suffering and you believe that this practice causes much unnecessary suffering. Instead of saying that gay marriage undermines "family values," say that it undermines *your* family's values, that it is against the teachings of your religion, etc. (Obviously some people will have an easier time with this transition than others. This is an important point to be explored later in Chapter 5.) Speaking in this way is honest, requires no false metaphysical commitments, and should make

---

[15] Haidt (Pg. 829) does make some suggestions that are similar in spirit to, though more moderate than, my own. He proposes, quite plausibly, that various forms of structured, active engagement with moral issues, learning to see moral issues from different viewpoints, may improve people's moral thinking, especially young people's.

discussions of moral matters much more fruitful and, at the very least, shorter. When someone makes a claim about how he feels; what he wants, values, or cares about; or what he is or is not willing to accept in a negotiation, there is, in the absence of realist interlocutors, nothing to dispute. When someone makes a putative statement of fact, there is often much to dispute, but in a purely factual discussion there is a decent chance that genuine evidence can be brought to bear on the issue, resolving it in favor of one party or another or, failing that, demonstrating to both parties that the evidence is inconclusive.[16] Even where differences of value and/or factual disagreements persist, such a mode of discourse is likely to lead to increased mutual understanding and less exasperating huffing and puffing.

The language of moral realism is sufficiently rich to provide a reasonable sounding justification for just about anything a society would actually want to do. Even terrorists (or "freedom fighters," depending on your point of view) can justify their actions in terms that sound eerily similar to those used by their victims in other contexts. Because there is no fact of the matter about what's right or wrong, no true moral theory, there is no neutral ground from which to sort out the putatively true moral claims from the ones that simply ring true to some people. Thus, the language of moral realism makes an excellent smoke screen for aggressive or otherwise anti-social behavior, a smoke screen that is so effective

---

[16] People on opposite sides of a practical issue will often have a hard time agreeing on the facts, but I suspect that this is because the "facts" in such cases are not really purely factual but rather mixed up with evaluative matters as well.

it can, and usually does, fool those who employ it. A nation with economic incentives to take over a neighbor's territory can claim that the neighboring territory *really* belongs to their nation and that they have a *right* to it. One might go so far as to say that nations *require* the language of moral realism to marshal popular support for aggressive actions. Has a military aggressor ever *not* claimed a moral right to carry out its plans? Has a nation ever been moved to war by leaders who said, "It would be good for us economically, and we can get away with it, so why not?" As Roger Fisher says in *Basic Negotiating Strategy* (1971, Pg. 110), "Governments are not like bank robbers. People ask 'What ought I do?' not 'What can I get away with?'" Arthur Schlesinger (1971) concurs: "National interest has a self-limiting factor. It cannot, unless transformed by an injection of moral righteousness, produce ideological crusades for unlimited objectives." As does Kenneth Thompson (1985, pg. 5): "Foreign policy tends to be articulated in moral terms, even in most authoritarian regimes."

As noted above, moral realists naturally interpret the actions and opinions of their (realist) opponents as products of a surreptitious disregard for morality. They see things this way because, to them, the moral truth that has been disregarded by their opponents is so self-evident that only one who willfully disregards morality's requirements could think and act as their opponents do. People who are evil, people who believe that their actions are wrong and carry on just the same, are native to Hollywood movies and children's books, but in real life (and good books and films) such characters are few and their influence is negligible. (See Section 5.2.3.) In the real world, the vast majority of avoidable

suffering is caused by people who think they have the moral truth on their side. Improving the world is not a matter of getting those who aim at The Bad to aim instead at The Good, nor is it a matter of clarifying the true nature of The Good so as to inform those people(s) who somehow managed to get the moral facts wrong. (In my experience, many moral philosophers take themselves to be foot soldiers in a campaign of the latter sort.) Rather, the challenge is to change the behavior of well-meaning people everywhere. It is a matter of understanding that the enemy is indeed "us" and not an evil and/or misguided "them." It's a matter of understanding why we are the way we are and how we can adapt our thought and behavior to life under the unnatural conditions we have created for ourselves.

Of course, we don't have to stop thinking and talking like moral realists just because moral realism is false. We could pretend that there is a fact of the matter about what's right and wrong, asking questions like, "Is capital punishment *wrong*?" and "Do the Palestinian people have a *right* to self-determination?" and then constructing arguments based on not-exactly-true premises so that we can then go on to say things like "No it's not!" and "Yes they do!" As noted above, this is the standard proposal made by those who reject moral realism. In Section 4.6 we will consider two specific proposals along these lines (quasi-realism and fictionalism) and the arguments given on their behalf. Another anti-realist approach is to construct a true moral theory either by changing the meanings of our moral terms or by developing a theory that embodies a particular moral sensibility rather than one that is supposed to be universal. We will discuss

these proposals (reforming analytic naturalism and constructivism) in Section 4.7.
Before moving on we will pause to explore an oft-made analogy between moral
properties and color properties that will set the stage for the upcoming critiques
of the anti-realist alternatives to revisionism.

4.5 The Analogy with Color

Many philosophers, both realist (e.g., McDowell, 1985) and anti-realist (e. g.,
Blackburn 1984, 1993a), have compared values to colors and other "secondary
qualities." Both camps agree that colors and morals are in some sense mind-
dependent. They disagree about whether this mind-dependence makes color
properties and moral properties less than fully real. In practical terms, however,
their disagreement amounts to very little. Both would have us carry on thinking
and talking about moral matters just as we do, and both compare ordinary moral
discourse to ordinary color discourse, which is taken to be perfectly acceptable:
Surely it would be silly to refuse to call the sky "blue" and fire trucks "red"
because of what scientists have taught us about the physical/psychological
bases of color. Surely it would be silly to replace our ordinary color talk with talk
about spectral reflectance properties or whatever mind-independent properties
"really" exist instead of colors. And if the analogy between colors and moral
values holds, then isn't it equally silly to change the way we think and talk about
morals?

No. The analogy between values and secondary qualities holds, but only imperfectly. The underlying metaphysics is roughly the same in both cases, but our circumstances with respect to these two aspects of experience are importantly different. To illustrate this point we can imagine a world in which the color-moral analogy holds better than it actually does.

Some people are color-blind. In most cases, they fail to perceive the colors red and green due to a deficiency in the visual subsystem that normally gives rise to our experience of these two colors. As a result, color-blind people see both blood and grass as colorless and therefore cannot distinguish between them on the basis of color. As it happens, many of the items that normal people take to have the same color have very different spectral reflectance properties.[17] Thus, it is perfectly possible that there could have been creatures—call them our "color twins"—who are like us in nearly every way, but who stand with respect to us and our visual systems as we stand to color-blind people and their visual systems. Just as we perceive blood and grass as differently colored, our color twins perceive many objects that look identically or similarly colored to us (e.g., tomatoes and a fire trucks) as clearly differently colored. We can imagine that this discrepancy in color perception is rather widespread. For every single color we see, they see two or more different colors.

Suppose we begin interacting with our color twins. We might not notice any difference between them and us at first, but eventually we would. At the outset we might attempt to help them get around our cities by, say, pointing to a

---

[17] See discussion of metamers in C. L. Hardin (1988).

tomato and saying, "When you see a light above the street that's *that* color, make sure you stop." After enough car accidents, however, we would recognize the importance of modifying our color discourse, at least for the purposes of communication between us and our color twins. Only ignorant and/or foolish people would persist in, say, making presentations with color-coded diagrams that only make sense to audience members like us and not to the color twins in the audience. At the same time we would hope that our color twins would present things to us using color-coded diagrams that we can easily interpret, and so on. Each of us would change the way we communicate in order to avoid misunderstandings and, in some cases, avert disaster.

Color twins are just a philosopher's fantasy, and since they don't exist it makes perfect sense for us to go on talking about colors just as we do. But this is not because the colors we see are the world's "true colors."[18] It's just that, for

---

[18] Some philosophers, of course, would disagree. The ones who object on linguistic grounds might claim that color property terms are covertly indexical and that "red" really means something like "red according to the color conventions we use," thus making "red" in the mouths of color twins either a misunderstood term of ours or a homophonic term of theirs. As a matter of psycholinguistic fact, I don't think that words like "red" are indexical in this way. But even if they are, the point still remains that it would be foolish to go on talking as we do about colors while trying to get along with our color twins. Some would disagree on pragmatic grounds: If it makes sense to say that tomatoes and fire trucks are the same color then it's *true* that they're the same color. I reject this conflation of what is

most of our purposes, they *might as well be* the world's true colors.[19] With the

exception of a few color-blind people,[20] there is broad consensus concerning the

intersubjectively verifiable aspects of color experience, and so it would indeed be

silly to talk as if our imagined color twins lived among us, even if doing so would

be more faithful to the true metaphysics of color.

The key question in pursuing the color-morality analogy is whether or not

the actual moral world is more like the actual color world or the imaginary world

---

true with what it makes sense to say, but no matter. The key point, once again,

is that in a world like the one in which our color twins are among us, it no longer

makes sense to speak as we do. And, finally, some philosophers might say that

our color talk is just correct, not because its terms implicitly refer to our color

standards and not for pragmatic reasons, but simply because we see the world

as it truly is while our color twins, if they existed, would see the world incorrectly.

In response to such philosophers I ask, much as I did with respect to morals in

Chapter 2, what makes our color claims true and those of our color twins false?

[19] Of course, there's the familiar worry about whether or not you and I really see

the same things when we look at the various things we call "red," "blue," etc., the

so-called "inverted spectrum" problem. This problem, however, is beside the

point for our purposes since we're only interested in intersubjectively verifiable

differences.

[20] There are other anomalies in color perception besides color-blindness. For an

example, see Ramachandran and Hubbard (2001) for a discussion of

synaesthesia.

in which we live among our color twins. After the discussion of moral psychology in Chapter 3 and the previous two sections in this chapter, the answer is clear. Our moral twins are all around us. Because of our varying cultural backgrounds, we humans see the moral world in a variety of ways.[21] When pro-life advocates claim that a fetus has a *right* to life and pro-choice advocates reply that a woman has a *right* to choose, they are like you and your color twin arguing over whether or not tomatoes and fire trucks are the same color. If there were a fact of the matter about whether or not these items are the same color, then at least one of you might have a good reason to stick to your guns. But there are no such facts. Likewise, there's no fact of the matter about whether a fetus' "right to life" outweighs a woman's "right to choose," and, if two combatants in this debate really understand what's going on, it's perfectly silly for them to try to influence one another with this sort of talk.

Thus, the analogy between moral properties and color properties speaks in favor of revisionism. But the analogy is, as noted above, imperfect. The color analogy suggests a kind of "moral relativism" in the popular sense of this term: If people who see moral matters differently from us want to rape, murder and

---

[21] Of course, there are many reasons why two people might see the moral world differently besides differences in cultural backgrounds. People have different natural temperaments and different experiences, even within the same culture or even within the same family. And some people even reason their way to different moral conclusions. The point here is simply that cultural differences alone are sufficient to produce "moral twins."

pillage, that's no worse than their using color-coded diagrams that make sense to them but not to us. I am most certainly *not* advocating this sort practical relativism. The analogy breaks down because of a crucial difference between the respective natures of color and value. Seeing the world within a particular chromatic framework allows one to make certain discriminations, but it does not provide one with *ends*.[22] If I, unlike you, see fire trucks and tomatoes as similarly colored, that implies nothing about how we ought (subjectively speaking) to differ in our behavior. In contrast, if I am a woman who values social and political independence, and you are a woman who shuns these things, we are likely to have very different practical agendas, and the meta-ethical facts give us no reason to abandon these agendas. Colors are (more or less) practically neutral. Values are not.

4.6 Faking It: Quasi-Realism and Fictionalism

4.6.1 Blackburn's Quasi-Realist Proposal

Most contemporary anti-realists would preserve the way we think and talk about moral matters except in our more reflective, philosophical moments. Perhaps the most prominent version of this position is Simon Blackburn's *quasi-realism* (1984,

---

[22] Perhaps it gives one certain aesthetic ends. Insofar as we are talking about aesthetics, the analogy continues to hold.

1993a). Blackburn, like myself and John Mackie, is a projectivist anti-realist with no inclinations toward nihilism. Much of Blackburn's efforts are aimed at attacking moral realists for their metaphysical extravagance, working out the details of his projectivist picture, and fending off accusations of nihilism. He also raises doubts about John Mackie's (1977) error theory on the grounds that Mackie takes apparently no interest in revising our moral practices.[23] This is all to the good. Where Blackburn and I part ways is, once again, over the issue of *revisionism*, particularly when it comes to the language of moral discourse. I think some serious changes are in order, whereas he not only denies that there are good reasons to make such changes, he thinks that it's more or less inevitable that we talk about moral matters the way we do. I will now argue against both of these claims, beginning with the second.

In the process of defending his quasi-realist descendant of expressivism against expressivism's old nemesis, the "Geach-Frege" problem (Geach, 1965), Blackburn (1984) sketches a formal semantics for quasi-realist moral discourse. He imagines a language much like English, but with evaluative operators instead of evaluative predicates. There's the "Hurrah!" operator "H!" for expressing positive attitudes and the "Boo!" operator for expressing negative attitudes. Thus, one might express one's negative attitude toward lying by saying, "B!(lying)" and one's approval of going to church on Sunday by saying "H!(going to church on Sunday)." By putting an attitude expression inside a pair of "|" one

_____

[23] That is, Mackie makes no proposals that are grounded in any obvious way in his *meta*-ethical views.

can refer to that attitude or to sensibilities that endorse that attitude, and one can

denote the conditional coupling of two attitudes by separating two attitude

expressions with a ";." Thus, one can express disdain for those who approve of

lying by saying "B!|H!(lying)|," and one can express one's approval of moral

sensibilities that reject lying only if they reject getting one's little brother to lie by

saying, "H!(|B!(lying)|;|B!(getting little brother to lie)|)." And with all of this come

rules of inference that parallel those of standard, deductive logic. Paralleling

modus ponens, one cannot coherently[24] endorse the claims "B!(lying)" and

"H!(|B!(lying)|;|B!(getting little brother to lie)|)" without also endorsing the claim

"B!(getting little brother to lie)." Blackburn adds an uncertainty operator "?" and

some other devices, all of which are designed to make his prototypical moral

language into "an instrument of serious, reflective, evaluative, practice, able to

express concern for improvements, clashes, implications, and coherence of

attitudes," (1984, pg. 195). Blackburn supposes that this language, properly

fleshed out, will look just like ordinary moral language, thus showing that a moral

discourse that is expressivist at its core can "earn the right" to the apparently

realist language we currently use, thus vindicating quasi-realism. He goes so far

as to declare that all revisionist projectivists will ultimately "find themselves

indulging in a practice that is apparently identical with moralizing [i.e.

moralizing$_1$]," (1993a, pg. 152).

---

[24] The coherence here is not logical coherence, but a kind of practical and/or

psychological coherence.

The great irony in this is that Blackburn has just given us a scaffold on which to construct a moral language that is *transparently anti-realist* and decidedly *not* identical to ordinary moralizing$_1$. His prototypical moral language does, as advertised, mirror ordinary moral$_1$ language in many ways. His language has the potential to be "an instrument of serious, reflective, evaluative practice" with all the bells and whistles: notions of consistency, uncertainty, second order attitudes, etc. But there is one feature that it does not, or need not, have, namely any suggestion of *moral realism*. One can "Boo!" and "Hurrah!" all day long and never sound like a realist. One can say "Boo!" to sensibilities that are uncertain about saying "Hurrah!" to sensibilities that endorse sensibilities that approve of deceiving people but not of deceiving cows, and so on, and so forth, and never come one bit closer to suggesting that there is any *fact of the matter* about what anyone ought to do. Blackburn's proto-language, rather than evolving into ordinary moral talk, could just as easily evolve into a sophisticated kind of moral$_2$ talk, one without the usual moral$_1$ overtones—in other words, exactly what we're looking for! Well, not *exactly*. All of this "Boo!" and "Hurrah!" stuff is a bit awkward for everyday use, although it could catch on. For now, we can stick with more familiar subjective expressions such as "I'm opposed to X," "I find X unacceptable," and "We are against X for the same reason we're opposed to Y, namely…" and so on. But Blackburn has the right idea, even if he doesn't know it.[25]

---

[25] Or didn't at the time. In a later work, Blackburn (1993b, pg. 379) acknowledges in passing that "something like a Boo-Hooray Language" could

But is it the right idea? The fact that Blackburn was wrong to think that revisionist projectivists[26] *have to* end up speaking like ordinary moralists$_1$ doesn't mean he was wrong to think that they should. How, then, should we decide whether or not to persist with ordinary moral discourse? Like nearly everyone else, Blackburn takes it for granted that it would be a good thing to hold on to our ordinary moral discourse. He views the revisionist challenge not as a progressive program with a worthy practical agenda but as a persnickety philosopher's attempt to spoil the party. He never weighs the costs and benefits of using ordinary, realist moral language against those of using a strictly moral$_2$ language of the kind he's inadvertently devised because he simply assumes that ordinary moral$_1$ discourse is the best sort of moral discourse to which we could hope to "earn the right." Is it?

Here my argument is little more than a reminder of what's already been said. In Chapter 3 I developed an empirically based, projectivist picture of moral psychology with some unanticipated features that make trouble for quasi-realism down the line, chief among them the fact that individuals absorb the local morality and then project it onto the world at large. In Section 4.3 I argued that this psychological tendency, encouraged by realist language, does not scale well, exacerbating conflict and causing much unnecessary strife and misunderstanding. In Section 4.4 I laid out the revisionist alternative to moral$_1$ discourse and explained why it is likely to avoid some of these problems. In

serve as an alternative anti-realist discourse.

[26] That is, projectivists who aren't nihilists.

Section 4.5 I explained how analogizing moral properties to color properties, properties that many take to be projected, ultimately speaks against the preservation of ordinary moral discourse. Blackburn (1984, pg. 171) describes the project of moral quasi-realism as one of showing that our ordinary moral talk is not so "diseased" as Mackie and others have claimed. I've argued that our ordinary moral talk is precisely "diseased," but with arguments that are nowhere anticipated by Blackburn. Thus, insofar as these arguments are compelling, the burden is on the quasi-realist to explain why ordinary moral language is worth preserving.

The argument made here against Blackburn also applies to realists like McDowell (1985) who wish to assimilate moral properties to secondary properties such as colors. The debate between Mackie (1977) and McDowell over whether moral properties are *real* and over whether moral judgments may be *true* is ultimately a debate about the nature of truth. Mackie holds out for something like a strict "correspondence" theory of truth while McDowell, Blackburn, and others favor a more inclusive view that will accommodate truths that fall short of being accurate characterizations of a mind-independent world. But now that the debate has shifted into the practical domain we can bypass these familiar and apparently endless debates about the nature of truth.

Mackie says that if it's not a mind-independent duck, then it's not really a duck. Blackburn says that if it's not a mind-independent duck, it may still be a quasi-duck if it manages to walk like one. McDowell and like-minded realists go further and say that if it walks like a duck, then it's a duck, mind-dependent or

not. My case against McDowell *et al*. is the same as my case against Blackburn: *It doesn't even walk like a duck*.[27] Speaking as if moral properties were out there in the world is, however natural, a foolish thing to do, and no adequate theory of truth will require us to speak foolishly.

4.6.2 Joyce's Fictionalist Proposal

Another anti-realist approach to moral practice is *fictionalism*. The fictionalist accepts that moral realism is false, but nevertheless advocates maintaining the fiction that moral realism is true, at least in most contexts. In *The Myth of Morality* (2001) Richard Joyce mounts what I take to be the most elaborate and compelling defense of fictionalism currently available, and for this reason I will take Joyce's view as representative of the fictionalist viewpoint.

Joyce recommends preserving our realist moral discourse because it is *useful*. More specifically, he believes it is an effective device for combating *akrasia* or weakness of the will. A general habit of thinking of some things as

---

[27] McDowell (1985) writes, "Can a projectivist claim that the position I have outlined is at best a notational variant, perhaps an inferior notational variant, of his own position? It would be inferior if, in eschewing the projectivist metaphysical framework, it obscured some important truth. But what truth would this be?" Answer: the terrible, horrible truth about morality *and what to do about it*. Realism and realist talk blinds us to the practical advantages of revisionism.

wrong and other things as *right* or *obligatory* is a good way to make sure that one is not lured into immorality$_2$ by life's temptations, even if one acknowledges in moments of meta-ethical reflection that nothing is *really* right, wrong, obligatory, etc. For example, one who would like to resist the temptation to shoplift will be that much more likely to succeed if he can tell himself that doing so would be *wrong*, much as one can motivate oneself to do *all fifty push-ups* by saying to oneself *"Must do all fifty push-ups!"* Joyce argues that while it's not *true* that you *must* do fifty push-ups, telling yourself that you must and pretending to believe it may help keep you in line. Moreover, Joyce argues that we can enjoy the benefits of pretending to believe in morality$_1$ without too much psychological strain. He compares this fictionalist practice to that of engaging with literary fictions. Reading novels and watching films requires no strange delusional beliefs. On the contrary, doing this is quite healthy because engagement with fictional narratives helps us regulate our emotions, gives us insight into our own lives, etc.

Joyce's case for fictionalism is surprisingly good. Nevertheless, I believe his account (1) underestimates the costs of fictionalism by ignoring the effects of realist thought and language on public life in a pluralistic world, (2) underestimates the benefits of revisionism by overlooking certain useful ways of thinking and talking that are available to revisionists, (3) underestimates the difficulties associated with maintaining the moral fiction in ordinary discourse, and

(4) overestimates the benefits of aiming for a smooth transition away from realism.[28]

First, in making his case for fictionalism, Joyce's focus is entirely on the function of morality under conditions of uniform moral values and not under conditions of moral disagreement or conflict. I've argued that the main problem with realist moral discourse is that it exacerbates conflict and generally makes it harder for people with conflicting values and/or interests to get along. Were our values more uniform, I'd be happy to be a quasi-realist or a fictionalist, much as I'm happy to retain ordinary color discourse even if color properties are, metaphysically speaking, on a par with moral properties. Like Blackburn, Joyce simply assumes that ordinary realist discourse is the way to go, provided that we can find a good rationale for keeping it in spite of its false presuppositions. He never considers the possibility that there might be serious practical costs associated with realist thought and language. Thus, here, too, my case against realist discourse is entirely unanticipated.

Second, Joyce never considers the possibility of employing a revisionist discourse of the kind I advocate. In accordance with the standard conflation of morality$_1$ and morality$_2$, Joyce assumes that one must either think in pseudo-realist terms or in egoistic terms. For example, at one point he implicitly assumes that a hospital "ethics committee" (shudder quotes are Joyce's) that eschews realist moral discourse would have to make their decisions concerning whether or not to conduct secret drug trials, surgical operations without consent,

---

[28] Points (3) and (4) are really extensions of point (1).

surreptitious acts of euthanasia, etc. from an *egoistic* standpoint (Pg. 220). But

this is simply false. Members of such a committee could care very deeply about

the well-being of experimental patients, the freedom of patients to guide their

medical care, and the desires of critically ill patients and their families. They

could talk about their values and act on them, all without any suggestion that

there is a fact of the matter about what they ought to do.

In another case, Joyce (Pp. 225-7) imagines three possible agents faced

with a choice of whether to join an unscrupulous friend in shoplifting. Agent 1 is

a moral realist who resists in the usual way, by thinking that it would be wrong to

shoplift. Agent 2 is a fictionalist who resists by telling himself and pretending to

believe that it would be wrong to shoplift.[29] Agent 3, the revisionist, has,

according to Joyce, "nothing to say" to himself and to his friend about why he

shouldn't shoplift and is therefore more likely than his realist and fictionalist

counterparts to give in to temptation. But why should Agent 3 have "nothing to

say?" Why can't he say, "It's against everything I stand for! I couldn't live with

myself as a thief!" Joyce (Pg. 226) claims that if Agent 3 were to give in to

temptation "it's not clear that he would feel anything more than surprise at

having done so." But why should he not experience the sickening feeling of

having betrayed his core values?

According to Joyce, all that stands between Agent 3 and a life of petty

thievery is a mere "habit," flimsy for having not been "moralized" (Pg. 227). But is

it lacking moralization$_1$ or moralization$_2$? One who has rejected morality$_1$ may

---

[29] As Joyce describes the case, the chances of getting caught are negligible.

nevertheless maintain a rich and complex moral$_2$ value structure of the sort we associate with realist moral commitments. Here I invoke something like David Lewis' (1989) Frankfurtian (Frankfurt, 1971) account of values as second-order desires. If while I'm busily working I find that hunger is impeding my all-important progress, I'd be happy to get a sandwich, but I'd also be happy to stop wanting one. I desire the sandwich, but I don't desire to desire the sandwich. In contrast, if I abhor racism, I'd be happy to see racism end, but I am *not* similarly happy to stop abhorring racism. It's very important to me that I desire against racism, and were I to stop I would consider myself diminished, to have lost a piece of myself. Thus, my desire against stealing may be *moralized$_2$* by being deeply ingrained, by my maintaining it not only as a desire, but as a desire that I desire to retain and act on. I believe that it is primarily this sort of psychological structure, and not the realist beliefs that often accompany it, that allows people to combat weakness of will. What's missing in children who are old enough to tell you the rules but not old enough to resist breaking them is not an (implicit) commitment to realism, but rather a fully developed sense of self, one with second order desires strong enough to overpower the first order desires that weak-willed children may already believe to be bad.

Third, Joyce suggests (Pg. 192-3, 196) that as fictionalists we can easily put aside our anti-realism when not engaged in philosophical reflection. If moral discourse were simply a way of bringing our thought and actions in line with a single moral system that we all accept then this might be true. But, once again, we live in a pluralistic world in which moral commitments are often challenged by

those who reject them, and while much moral talk is among parties that share the same moral viewpoint, the moral talk that matters most may well be the moral talk among people attempting to live together in spite of their moral disagreements. Recall the analogy with color. In a world like ours in which everyone sees colors the same way, it's easy to pretend that tomatoes and stop lights *really are* the same color, but this would not be so were we living among our color twins, much as we presently live among our moral twins. Fictionalism, when it works, is a kind of benign propaganda that one applies to oneself, but in the context of conflict with other fictionalists it becomes a kind of pointless, confusing, and potentially inflammatory propaganda applied to others.

The problem with fictionalist discourse extends beyond cases of conflict to cases of genuine inquiry. In such contexts we normally talk about first-order moral issues as if we were in search of facts. People ask "Is abortion wrong?" "Do murderers have the right to live?" and so on. It would be a bit odd for an anti-realist to use terms like "right" and "wrong" in such non-asserted contexts. One could do it, but what would be the point? In such cases we're not fighting weakness of the will, which is the main reason that Joyce gives for adopting fictionalism.[30]

---

[30] Joyce considers such cases and suggests (Pg. 219) that thinking in moral terms in such contexts would encourage moral thought in cases in which weakness of will is an issue. But, as I argue below, the revisionist need not forgo the benefits fictionalism affords in such cases.

Fourth, Joyce suggests that even if the ultimate goal is revisionism, fictionalism is a nicer option for the near future because it will make for a smoother transition from realism. I take the opposite view. Perhaps someday we could be fictionalists if that proved convenient, but what we need now is as clean a break as possible with the pernicious habits of realist thought and language. Consider a comparison between realist language in moral discourse and teleological language in discussions of evolution. Students of evolution must learn that natural selection is not a forward-looking process. To help this lesson take hold, beginning students are urged not to speak in teleological terms. Experts, however, are free to enjoy the convenience of saying things like "Giraffes developed long necks in order to reach more food," because they know better and everyone knows they know better. Likewise, I suggest that we should revert to realist language for the sake of convenience only after anti-realism may be taken for granted.

Finally, even if one is convinced that revisionism is a generally better policy than fictionalism, one might think nevertheless that fictionalism can be a useful tool for fighting *akrasia* and regret the thoroughgoing revisionist's categorical rejection of it for that reason. Fair enough. If fictionalism proves useful in navigating one's way through the stormy waters of one's personal life, and if one can use it without causing damage elsewhere, then why not keep it around? If telling yourself that it would be *wrong* to have an affair helps you remain faithful to your spouse, then, by all means go ahead. You can even pretend that marital infidelity will send you *straight to Hell* if that suits you. Moral

fictions maintained in the privacy of the home among consenting adults are fine with me.

More seriously, it might turn out that a healthy dose of moral fiction is useful or even necessary for raising moral$_2$ children. According to the account of moral psychology given in Chapter 3, moral experience tends to have a realist phenomenology and as such inclines people toward a realist interpretation of moral rules. Children who lack the intellectual sophistication to distinguish between morality$_1$ and morality$_2$ may have no choice but to be (implicit) realists insofar as they are moral$_2$. Thus, it's quite likely that even if one speaks to one's child using the sort of transparently anti-realist language I favor for adults the child will nevertheless interpret this language realistically. Or worse, such language may confuse the child or fail to be understood at all and thus prevent him or her from learning the important moral$_2$ lessons she needs to learn. Thus, using revisionist language with young children could be pointless or even dangerous. Is it?

This seems to me unlikely. Recall from Chapter 3 the central role of emotional communication in moral development. If the emotions are what really matter, then an emotionally charged, "*Don't hit your sister!*" or "*We don't tolerate hitting in this house!*" is likely to be just as effective as a, "*Hitting is wrong!*" A child who is exposed to such admonishments will very likely have an implicitly realist picture of moral life in spite of what she hears, but speaking this way won't be pointless insofar as it will facilitate the transition to revisionism later on. Whether or not moral$_1$ language is necessary for the raising of moral$_2$ children is

an empirical question, and if it turns out to be necessary then we can indulge our children in the realist fiction until they're old enough to know the truth about morality, Santa Claus, and whatever else.

4.7 Remaking It: Constructivism and Reforming Analytic Naturalism

4.7.1 Constructivism

I will not offer a thorough critique of ethical constructivism, but I would like to make a few remarks concerning this view since many people take it to be a viable *anti-realist* approach to ethics. The archetypal constructivist is John Rawls, and I focus my brief comments on his work. Elsewhere I offer a more elaborate critique of the Rawlsian account of justice (Greene and Baron, 2001).

Constructivists sound a lot like realists. Rawls' *magnum opus* is entitled *A Theory of Justice* (1971) rather than, say, "My Conception of Justice" or "The Hyper-Egalitarian Western Liberal Conception of Justice" or something like that. Most of the book reads like an explication of what justice *really is*, and if that's what constructivists such as Rawls are really up to then their arguments are subject to the skeptical doubts raised against realism in Chapter 2. However, when pressed, Rawls denies that he is working out the true principles of justice and insists instead that he is uncovering the underlying principles behind one

conception of justice that many people, especially participants in Western-style democracies, implicitly share:

> This rendering of objectivity implies that, rather than think of the principles of justice as true, it is better to say that they are principles of justice most reasonable for us given our conception of persons as free and equal, and fully cooperating members of a democratic society (Rawls, 1980).

On the one hand, If Rawls' conception of justice were thoroughly idiosyncratic, then it would make no sense at all for him to label some institutions etc. "just" and others "unjust" when these labels indicate nothing more than compliance or failure to comply with principles that accurately reflect *his* sense of justice and no one else's. If, on the other hand, Rawls' conception of justice is implicit in nearly everyone's moral and political thought, then it would arguably be quite appropriate to apply the labels "just" or "unjust" to institutions etc. that embody or fail to embody this more or less universal sense of justice. The question, then, is whether or not Rawls' conception of justice is sufficiently universal to warrant the adoption of apparently realist language for evaluative public discourse. Is it?

This is extremely unlikely. Rawls' conception of justice is highly rationalistic, and from Chapter 3 we know that rationalism as an account of moral psychology, as an account of people's sense of right and wrong, is highly suspect. It's unlikely that Rawls' principles accurately reflect anyone's intuitive sense of justice *including John Rawls*. Rawls' theory of justice is a made-to-

order illustration of the sort of post-hoc rationalization that Haidt's social

intuitionist model finds in nearly all moral reasoning[31] and that casuistic moral

philosophers so readily dismantle. But even if Rawls grants that his theory is an

attempt at post-hoc rationalization of his more basic moral intuitions,[32] why

---

[31] At a crucial point in Rawls' discussion (1971, pp. 150-161), the normative

upshot of his view depends on which principle for choice under uncertainty the

individuals behind the veil of ignorance would use in ranking, from a self-

interested point of view, the various sorts of societies they might create for

themselves while not knowing which positions in society they would occupy.

Rawls decides, for reasons that are hardly compelling, that they would use a

"maximin" principle, one that gives absolute priority to increasing the position of

those worst off in society, a principle of maximum risk aversion. According to this

principle, it would be, in purely self-interested terms, more rational to lottery

oneself into a society in which everyone is very poorly off than into a society in

which everyone is fabulously well off except for one person who is slightly less

well off than the people in the first society. Would anyone without an ulterior

motive really take this decision principle seriously? (Remember that it's

supposed to a be a principle for purely *self-interested* choice.) Would Rawls do

so if it didn't give him the sort of egalitarian outcome that he wants out of a theory

of justice?

[32] Which he essentially does. See the discussion of the method of reflective

equilibrium in Rawls (1971, pp. 46-53).

should one think that there even *is* a snappy set of principles that captures his or anyone else's sense of justice?

Remember the trolley problem. Most people say that it's okay to hit the switch so that the trolley runs over one person instead of five. Most people also say that it's not okay to push a large stranger in front of the trolley in order to save the five people on the tracks. The second action offends our sense of justice. The first one does not. Once again, there is good evidence to suggest that the second action, unlike the first, offends our sense of justice because we have a negative emotional response to it (Greene *et al.*, 2001). If Haidt is right in claiming that our emotions come first with our moral judgments and rationalizations thereof following after, then why should we expect our emotional responses to parallel a nice, neat theory of justice that will apply not only to situations like those that arose during the time when our social-emotional dispositions evolved, but also to the highly complex and biologically unprecedented circumstances of modern life? There's a reason why no one has ever formulated an intuitively satisfying, direct moral theory.[33]

But it's worse than that. The preceding account of moral psychology maintains that the human moral instinct, like the human language instinct, is universal in form but local in content. In other words, there are likely to be many different conceptions of justice, just as there are many different languages. These various conceptions of justice may overlap in content. It may turn out, for example, that the trolley problem intuitions mentioned above hold cross-culturally

---

[33] See Section 2.3.1 on direct v. indirect moral theories.

and that other intuitions do so as well. Nevertheless, it is highly unlikely that there will be anything close to cross-cultural consensus on general principles for organizing the basic structure of society.[34]

Perhaps that's expecting too much. Might there not be a conception of justice that is common to, say, all Americans? Again, not likely, given the fact that Americans include people from many different cultures. But even if we limited ourselves to, say, white, third-generation U.S. citizens, we wouldn't get what we're after. George Lakoff (1996) makes a strong case that conservatives and liberals have very different moral worldviews, worldviews that are unlikely to be compatible with a single moral construction. Hoping for nationwide consensus seems incredibly optimistic when Rawls' conception of justice won't even fly with certain members of his own department (Nozick, 1974). If Rawls were offering us an account of the truth about justice rather than a common sense of justice, that would be a different story, but I'll be darned if deep down Robert Nozick's sense of justice is captured by Rawls' hyper-egalitarian

---

[34] I've argued elsewhere (Greene *et al*., 2001) that the intuitive difference between the trolley and footbridge cases turns on an intuitive psychological distinction between what I call "personal" and "impersonal" harm. I doubt that this psychological distinction is likely to map onto any sort of normative distinction that we would want to endorse as part of a larger normative theory because such a principle would end up being something like, "It's wrong to kill one person in order to save several others when killing the one person involves an 'up close and personal' interaction and killing the others does not."

difference principle. In the end, Rawlsian constructivism makes the same mistake as Blackburn's quasi-realism and Joyce's fictionalism. It assumes that there is sufficient uniformity in people's underlying moral outlooks to warrant speaking as if there is a fact of the matter about what's "right" or "wrong," "just" or "unjust."

4.7.2 Reforming Analytic Naturalism

In Chapter 2 I argued against realist versions of analytic naturalism, the view according to which there are substantive moral principles that are true by dint of meaning. In Section 4.1.4 I applied some of the psychological lessons from Chapter 3 to the case of analytic naturalism, offering a general account of these theories that explains why they are bound to fail. Here I tie up a loose end from the aforementioned discussion in Chapter 2. The analytic naturalist, after realizing that there are no analyses of moral terms that yield useful moral principles, might go on to recommend, as a practical anti-realist proposal, that we change what we mean by our moral terms so as to generate, by an act of linguistic will, moral principles that are true by dint of meaning.

This would be a silly thing to do, kind of like making yourself "rich" by using the word "rich" to mean what we normally mean by "poor." For one thing, since this is supposed to be a practical proposal, it's not at all clear how one goes about deliberately changing the way people use words, especially everyday

words like "right" and "wrong." At best, the change would be incomplete, which would only lead to more confusion. Second, the analytic naturalist's moral principles are always *indirect*[35] which means that they never offer any real guidance. If we were to redefine "wrong," as, say, "that which we would desire against if we were fully informed and fully rational," then we would end up sounding a lot like moral skeptics, thus defeating whatever purpose there may have been in preserving an illusion of realism. Whenever someone asks whether some action is "wrong" we'd have to say that we really don't know since we don't know what our fully informed and fully rational counterparts would say or do. Of course, we might think that our own moral intuitions offer some guidance in this regard, but for practical purposes they're no help at all. If all concerned parties have the same intuitions, then there's no practical need to appeal to moral principles. If people have different moral intuitions, such principles offer no help in adjudicating among them. And which reforming definitions would we choose? If we had a basis for making *that* decision we'd be well ahead of where we are.

Reforming analytic naturalism is a non-starter, a half-baked notion of an inkling of a thought about how we might make true moral principles out of nothing. The more reasonable hope in the neighborhood of analytic naturalism is that we might someday come to a consensus about how we ought to organize our public and private lives and that maybe, if some clever philosophers manage to get their hands on this special moral theory before it's taken hold, we could

---

[35] This is, once again, because direct, general analytic principles fall prey to Moore's open question argument.

attempt to institutionalize it by building it into our language. I'm somewhat optimistic about the long-term possibility for moral convergence. Like many philosophers, I'm not immune to the thought that if everyone were sufficiently well-educated and clear-headed they would see moral matters *just like me*. And perhaps someday they will. But trying to foist one's substantive moral views on the rest of the world by changing the meanings of words would be foolish even if it were possible. Language does matter, and to some extent you can change the way people think by changing the language they use, but you can't create substantive moral truths by rewriting the dictionary.

4.8 "Just Causes" and the Rhetoric of Realism

Above I allowed that realist moral talk may be necessary for raising moral$_2$ children. This could be true without making much trouble for the revisionist program, however, as we would simply allow or even encourage realist dialogue with those who are too young to handle the meta-ethical truth. I also allowed that realist talk may be useful in certain private contexts and therefore made a similar caveat pertaining to those cases. A more serious problem arises for the revisionist plan if there are good reasons for maintaining realist discourse that apply to public discussions among adults. Can such a case be made?

Realist talk has a nice ring to it. Saying that something is "just plain *wrong*!" packs a lot more rhetorical punch than saying that it's "something we're

very much against!" I trust that revisionists will find creative ways to boost their rhetorical amplitude, but the possibility remains that there will never be any talk quite so effective as good-old-fashioned realist rhetoric. Suppose that's the case. Should we deny ourselves access to the most effective tools available for promoting our substantive $moral_2$ ends? Could the substantive ends of, say, the human rights movement really be adequately served without talk of human *rights*? Earlier I suggested that leaders may need realist rhetoric in order to garner support for nasty, aggressive actions. What about noble, progressive actions? Will soldiers fighting "just wars" stand by their posts as death draws near if they haven't been taught that *duty* demands it?[36]

These are ultimately empirical questions, but I suspect (surprise!) that these sorts of behaviors will persist in the absence of realist talk or thought. I appeal once again to the social-intuitionist model described in Chapter 3. $Morality_2$ is primary, while $morality_1$ is essentially a post-hoc interpretation of moral experience. We do the things we do not because of our meta-ethical beliefs, but because of who we are as individuals and as a species. This is not to say that $morality_1$ exerts no force on our behavior. If it didn't, it wouldn't be worth jettisoning. The point is simply that it's not the basis for our *pro-social* tendencies.

Still, the worry was not that we can't get by in general without realist talk, but that sometimes we may really need it. Here I make the same sort of caveat as before. If the only way to get the local tribesmen to stop mutilating the

---

[36] See Foot (1974).

genitals of the local tribeswomen is to make a big harangue about "human rights," then, by all means, do what you've got to do. But allowing for this sort of exception is consistent with the position that realist language is generally detrimental and generally worth avoiding.

4.9 Summary and Preview

Let us return to the question with which we began this chapter: If moral realism is false, what are we to do? My answer is simple enough: Stop thinking and talking like moral realists. In Section 4.1 I did some ground-clearing, arguing that my "error theory" is ultimately optimistic and that rejecting moral realism doesn't mean rejecting morality$_2$. I also clarified my subjective use of evaluative terms and gave a psychologically-based account of the failure of analytic naturalism. In Section 4.2 I laid out my criteria for evaluating the various practical approaches to anti-realism, arguing that we should accept or reject anti-realist proposals based on how well they further our core values. In Section 4.3 I laid out the core argument of this chapter. Drawing on the picture of moral psychology developed in Chapter 3, I argued that our natural tendency toward realist thought and talk has been making trouble ever since our circumstances diverged from those in which our moral psychology evolved. We're using Stone Age moral psychology in the Nuclear Age, and the costs of doing so continue to mount as the world gets smaller. In Section 4.4 I outlined my revisionist proposal, explaining how

abandoning the language of moral realism is likely to avoid these sorts of problems. In Section 4.5 I discussed the familiar analogy between moral properties and color properties, arguing that the analogy, properly drawn, actually favors revisionism over its alternatives. In Section 4.6 I considered two related alternatives to revisionism, quasi-realism and fictionalism, explaining why neither of them is likely to further our ends given the nature of moral psychology and our present circumstances. Similar arguments were made against Rawlsian constructivism and reforming analytic naturalism in Section 4.7. Finally, In Section 4.8 I considered the case for preserving realist talk on the grounds that it is valuable as a means of public persuasion in the service of moral$_2$ ends and concluded, drawing once again on my account of moral psychology, that such talk is not worth keeping in general, although possibly worth keeping in certain special circumstances.

Thus, in this chapter I've made a case for revisionism over its anti-realist alternatives. In the next chapter I will assume that revisionism has won our allegiance and attempt to flesh out my practical proposal by addressing some more fine-grained questions concerning the positive revisionist program: What role is there for ethics as a domain of philosophical inquiry? What sort of first-order moral positions are we, as revisionists, likely to adopt? How exactly should we attempt to advance our revisionist program? What obstacles are we likely to face and how should we attempt to overcome them?

Chapter 5

Revisionism in Practice

We've concluded that moral realism is false and that, given the way our minds work and the state of the world, we'd best stop thinking and talking as if moral realism were true. In the last chapter we considered, in a preliminary way, what this means in practice, but many important practical issues remain to be explored. This final chapter is devoted to exploring them.

A warning to the reader: The main philosophical work of this essay is behind us, and much of what follows is highly speculative—that is, even more speculative than the material covered so far—and, I confess, a bit preachy at times. That is not to say that the issues addressed below are unimportant. On the contrary, I believe they are of paramount importance. I simply wish to acknowledge that critical readers will find in what follows even more than usual with which to quibble. My aim, however, in presenting this material is not to establish firm conclusions but rather to raise some important questions and suggest a set of plausible answers.

5.1 Revisionism and the Future of Moral Philosophy

Many, if not most, moral philosophers aim to discover, or help discover, first-order moral truths. But if there are no first-order moral truths, then what, if anything, are moral philosophers supposed to do? While many projects in moral philosophy lose their interest once moral realism has been rejected, there remains nevertheless plenty of good work for moral philosophers, although not as much, perhaps, as one might have hoped. We turn now to the dos and don'ts of revised moral philosophy.

5.1.1 Normative Ethics and the Organization of Moral Intuitions

As normative ethicists we are interested in two things: answers to questions about right and wrong and the proper justifications for those answers. John Rawls (1971, pp. 46-53) described and canonized the standard dialectical approach to answering first-order moral questions as a method of "reflective equilibrium." We begin with a wide variety of intuitions about particular moral questions, some of which enjoy our confidence, some of which do not. The ethicist aims to develop moral principles that organize the intuitions in which we have the most confidence and then, if possible, use those principles to inform our thinking about the cases in which our intuitions are less clear. For realists, this sort of activity has two purposes. First, assuming that our moral intuitions

generally reflect the moral truth, organizing our intuitions with normative principles deepens our understanding of moral reality by revealing its underlying structure. Second, with approximations of the true moral principles in hand, we may adjudicate difficult cases by appeal to those principles, thus expanding our knowledge of right and wrong and correcting misguided or muddled intuitions.

Clearly, the realist version of ethics by reflective equilibrium is ruled out by revisionism. But what about a parallel anti-realist mode of inquiry by which philosophers, rather than revealing the underlying structure of objective moral reality, reveal the underlying structure of our subjective moral values? This program is perfectly acceptable so far as anti-realism is concerned, but I maintain that it is nevertheless unlikely to be of much use. The problem, as ever, stems from moral psychology. According to the social-intuitionist model, our moral judgments are driven primarily by emotional responses. For the most part, we don't reason our way to particular moral conclusions. Rather, we arrive at our particular moral conclusions via emotionally-based intuitions, and then, if necessary, attempt to make rational sense of those intuitive conclusions after the fact. But these attempts never coalesce into a coherent moral theory because the intuitions that constrain them are not so well behaved.

But why aren't they well-behaved?   Couldn't the process go something like this?: We learn, in bits and pieces, a coherent set of explicit principles early on, but after a while our knowledge of these principles is transformed from explicit, propositional knowledge of moral principles into a set of emotional intuitions. Because our emotional intuitions are the psychological descendents of

an explicitly held, coherent set of principles, those intuitions will be well-behaved. Moreover, there's no reason why we can't raise to the surface the moral principles that underlie our intuitions by watching our intuitions in action, much in the way that an experienced typist can reconstruct the layout of the keyboard by watching his fingers as he types in the air.

It's a nice thought, but the evidence speaks against it. The emotions/intuitions appear to come before the explicit theorizing in the development of both individuals and species. In humans, the ability to learn and manipulate explicit moral principles develops later than the social and emotional instincts that form the basis of our moral intuitions, and chimpanzees exhibit social behaviors and instincts that strikingly resemble our own without the benefit of explicit moral principles. For us, as well as our primate ancestors, the emotional dog wags the rational tail.

If that's right, our moral intuitions will be organizable by moral principles only to the extent that natural selection and the blind cultural processes[1] that shaped our social instincts happened to organize them for us, and it unlikely that they have. Put aside, for a moment, the profound influence of culture on moral intuition and consider the design pressures imposed on our social instincts by natural selection. Natural selection designed us to spread our genes in environments like the ones in which our ancestors evolved. Why should we think

---

[1] Not all cultural processes that shape moral intuitions are blind, but the ones that aren't blind, i.e. deliberate attempts to make moral rules have, by hypothesis, been ruled out as the primary shapers of our moral intuitions.

that the collection of instincts that nature gave us for this purpose should fit together into a nice, normative package? Take, for example, some standard intuitions concerning infanticide, abortion, and the killing of animals for food. (See Singer, 1979.) (1) Many people think that for a pair of parents to allow their infant to die, say, by withdrawing life-support mechanisms, is wrong. This makes sense from an evolutionary/psychological perspective because allowing one's child to die is not good for one's genes, at least under most circumstances.[2] (2) Many people think that abortion is permissible, though many disagree. This intuition, as well as the controversy that surrounds it, makes evolutionary/psychological sense. Abortion was not an option for our hominid ancestors,[3] thus explaining why people's intuitions about abortion vary. However, abortion shares many features with choosing death for one's infant (not to mention ordinary killing), which explains why many people develop strong intuitions against it. (3) Most people think that eating fairly intelligent animals such as pigs is morally acceptable. This, too, makes sense from an evolutionary perspective. Eating animals such as pigs is generally good for the survival of one's genes. (1)-(3) make sense from an evolutionary/psychological perspective,

---

[2] In some cases, however, ending one's child life can be biologically advantageous, particularly when the resources required to care for the infant could be used more effectively for the raising of other children. See Hausfater and Hrdy (1984).

[3] Our ancestors could certainly terminate unwanted pregnancies, but not without seriously risking the mother's life.

but at the same time these intuitions are very hard to make sense of using normative principles. If one argues that it is acceptable to kill a fetus but not acceptable to allow an infant to die on the grounds that fetuses and infants differ in their levels of intelligence, sentience, consciousness, etc., then one has to explain why it's acceptable to kill pigs when pigs are more like infants than fetuses in terms of these features. If one argues that it's acceptable to kill a pig but not allow an infant to die on the grounds that an infant has the potential to develop into a full-blown person while a pig does not, then one has to explain why it's acceptable to kill a fetus, which, like an infant, has the potential to develop into a full-blown person. As far as I know, no one has come up with a set of principles that does justice to these intuitions without running into intuitive trouble elsewhere.

I maintain, once again, that this sort of moral theorizing fails because our intuitions do not reflect a coherent set of moral truths and were not designed by natural selection or anything else to behave as if they were.   And note that this is the case for a *single person's* intuitions. Troubles only multiply when one must reconcile the conflicting intuitions of many different people. Thus, while anti-realism does not rule out the possibility of reinventing normative ethics as an attempt to organize our moral intuitions and values, it's not likely to work. If you want to make sense of your moral sense, turn to biology, psychology, and sociology—not normative ethics.

5.1.2 The Study of "Thick" Moral Concepts

With few exceptions,[4] moral philosophers have given up on producing first-order moral theories, theories that purport to tell us how to live, what actions are right or wrong and why.[5] This is not because moral philosophers wouldn't like to produce such theories, but rather because producing such theories, at least ones that square with most people's first-order moral intuitions, has proven very difficult (for the reasons given above). As a result, philosophers have redirected their efforts. Like me, many philosophers have turned toward meta-ethics and, to a lesser extent, the empirical investigation of morality as a psychological and sociological phenomenon. But these days, moral philosophy is increasingly dominated by a style of inquiry that is not empirical (at least not in any straightforward way), not so general as meta-ethics, but very abstract and, unlike normative or applied ethics, rather removed from practical moral issues. Much, but not all, of this work goes under the headings of "virtue ethics" and "practical reason" theory. Such discussions typically center around one or more "thick" moral concepts, concepts that have both descriptive and evaluative components such as *integrity*, *reason*, *flourishing*, *forgiveness*, and so on. When working in this vein, one typically begins one's discussion by making a number of subtle interpretive distinctions concerning one or more of these concepts. One uses a

---

[4] The main exceptions are utilitarians such as Singer (1972, 1975, 1979, 1995).

[5] Using the terminology employed in Chapter 2, these are *direct* and *general* theories.

variety of examples to illustrate these distinctions, some made up, some drawn from literature or popular culture. Then one cordons off one or more of these overhauled concepts for further scrutiny, acknowledging that the ones left aside are legitimate, though perhaps not as central as the ones selected for further discussion. One then makes various suggestions concerning the interrelations among these various concepts, explaining how *integrity* "in this sense" is a precondition for *flourishing* "in this sense," or how *forgiveness* has both a cognitive and emotive component, contrary to what so-and-so said, and so on. Often the discussion culminates in a rather detailed theoretical construction that binds together the concepts under consideration, something like: "To *flourish* is to *successfully* exercise one's *virtues* in the pursuit of long-term *goals* in the context of a *human life*." No stands are taken on any real-world moral issues, but, one hopes, the discussion has "shed light" on some of the concepts and terms that people use when they talk about real-world moral issues.

The pertinent question for us is whether or not there is a place for this sort of theorizing given our rejection of moral realism and any pretense thereof. Theories centered around "thin" moral concepts, purely evaluative ones like *right* and *wrong*, are clearly suspect, but there is, perhaps, a better case to be made on behalf of moral theories that trade in thick moral concepts. After all, are we saying that there's no such thing as *flourishing*? Or that there are no moral *reasons* for action? The dual nature of these concepts as both evaluative and descriptive[6] makes them difficult for revisionists to dismiss entirely. Were we so

---

[6] By "descriptive" I mean "value-neutral" as I used the term in Chapter 2.

inclined, we could attempt to sort out the descriptive and evaluative components within of each of these concepts, put our stamp of approval on the descriptive components, and throw out the evaluative components, at least for the purposes of factual discourse. In considering the theoretical claim above, for example, we could attempt to subtract out the evaluative components in *flourish*, *success*, *virtue*, *goal*, *human*, and *life*, and see if we're not left with some kind of true psychological or sociological claim. Or we could leave the evaluative components in and see if what we have is not in fact a purely conceptual moral truth such as, "It is *wrong* to not do what one *ought* to do." What we can't do is interpret some of these terms in a purely descriptive way and others in a partially or wholly evaluative way and hope to be left with something *true*. For example, suppose we let *flourish* retain its evaluative meaning and let the rest of the terms go completely descriptive. What we'll end up with is a claim asserting the necessary[7] equivalence of the evaluative property of *flourishing* on the one hand with some complicated descriptive property on the other. It will follow that whatever has this descriptive property is *good*, since to be an instance of *flourishing* is, among other things, to be something good. Thus we will have deduced a moral fact of the sort that we rejected in Chapter 2 and thus produced

---

[7] If the equivalence isn't necessary then the whole claim is empirical and one is essentially doing armchair science—that is, unless one thinks that there are non-empirical ways to discover contingent facts, which is unlikely given our background assumption that all the facts with which we are concerned supervene on lower-level facts of the sort investigated by science.

a *reductio* of the claim in question. It seems that such theoretical claims, insofar as they are intended to be true, will either be armchair empirical claims, conceptual truths concerning concepts that are never fully instantiated, or claims that are simply not true.

The foregoing is, of course, just a quick treatment using one example as a stand-in for a large research program, a program about which there is certainly more to be said than I will say here. But, at the same time, what more there is to say is, I suspect, of little consequence. After all, it shouldn't come as too much of a surprise that this apparently realist branch of moral philosophy serves little purpose once moral realism has been rejected, and, indeed, many of its proponents would agree. But is it completely pointless without realism? That will depend on what one hopes to achieve. If one's aim is to survey the moral landscape as a preamble to answering questions about right and wrong, then it's hard to see the point. If, on the other hand, one's aim is to toss around some ideas in a way that "sheds light" on our values and/or various aspects of the human experience, then it's simply a question of how well those ends are served by this sort of discussion. I suggest that the language of thick moral concepts only confuses matters. If one's ultimate aim is to make a claim about human nature, then why not make this plain? Instead of offering a "virtue-theoretic account of flourishing," why not just say that, as a matter of fact, living a certain kind of life is likely to make you happier, healthier, etc.? The reason why not, I suspect, is that it makes one's theory sound like nothing more than a bit of armchair social science. But if that's what it is, then that's what it is, and a bit of

armchair speculation isn't necessarily such a bad thing. Or perhaps the aim is

not to describe human nature, but to recommend a way of life. If that's the case,

then why not present one's recommendation as such? Why not simply say, "I

suggest we live like *this* rather than like *that*?"

5.1.3 What Ethicists Can Do, And do Well

Things aren't looking good for ethicists. Constructing first-order moral theories is

out. Tossing around thick moral concepts probably isn't very useful.   What's

left? Meta-ethics? Too late, I'm afraid. This was an exciting area of research

until Chapters 2 and 3 ended it. I say this in jest, of course, but not entirely. I've

no illusions about the fact that meta-ethical inquiry will go on whether I like it or

not, but insofar as my arguments are correct, the truth about meta-ethics is pretty

simple, and there's not much more to say about it.   …Is all hope lost?

Applied ethics to the rescue. Organizing our moral intuitions into a single

theory may be impossible, but at the same time we're not going to let our moral

intuitions, and, more importantly, our moral behavior, slop all over the place.

More specifically, we will continue to care about *consistency*, even if we are

unable to achieve consistency across a wide range of moral issues.[8] But what

do we mean by "consistency?" What sort of behavior is *in*consistent? Suppose

---

[8] That is, without subverting some of our moral intuitions.

we save the whales on Monday and then slaughter them on Tuesday. It's

certainly *possible* to do this. In what sense is this policy inconsistent?

While this behavior and the rule that guides it are not *logically*

inconsistent, they may nevertheless be inconsistent in one or both of the

following ways. First, this policy appears to involve the pursuit of incompatible

ends. If we suppose, as seems plausible, that our motivation in saving the

whales on Monday is to preserve their lives, not just for Monday, but for Tuesday

and beyond, then killing the whales on Tuesday appears to thwart Monday's

conservation efforts. Thus, the inconsistency follows from the introduction of a

further assumption about our goals.

In this example, the inconsistency is apparent at the outset, but we may

be engaged in other patterns of behavior that are similarly inconsistent, though

less obviously so. Take, for example, the devotion of immense resources to

saving individuals in trouble rather than to preventative measures that would

prevent larger numbers of other individuals from getting in trouble in the future.

For example, a mining corporation might spend millions of dollars and risk

several people's lives to rescue a single worker from a fallen mine shaft when it

could use that money to set up safety precautions that would save several

people's lives without putting any rescue workers at risk.[9] A similar issue has

arisen concerning the use of certain vaccines. Some vaccines have a very small

chance of infecting the people who take them with the diseases these vaccines

are designed to prevent, and some people oppose the use of such vaccines for

---

[9] This and next example are from Baron (1998).

this reason, even though the vaccines tend to save many more lives than they claim. To take another example, some have opposed clinical drug trials in which people who might otherwise receive a potentially life-saving drug receive placebo pills, even though the patients involved have no other access to medical care and the future of such potentially life-saving drugs depends on their being validated in clinical trials using placebos. Whether or not the policies described above are really inconsistent depends, as in the case of the whales, on what our goals are: Are we simply trying to save the most lives, or are there other considerations that matter as well?

To address these issues, ethicists can (1) identify cases in which our behavior is potentially inconsistent in this way, (2) examine the facts in order to determine whether or not there is, indeed, the potential for such inconsistency, and (3) examine our goals (and the values that underlie them) in order to determine whether or not our behavior is ultimately counterproductive.

The other kind of inconsistency that ethicists may attempt to root out is better understood as *arbitrariness*. Suppose that the aforementioned whale policy isn't some sort of a compromise plan between conservationist and business interests. Suppose instead that it is a policy enacted by a group of people who are not terribly concerned with the future of whale-kind, but who are nevertheless whole-heartedly in favor of saving whales on Mondays and whole-heartedly in favor of destroying them on Tuesdays. In that case, their behavior is not at all inconsistent in the sense of being self-defeating. It is, however, very strange. More specifically, it's very *arbitrary*. The fact that it's Tuesday rather

than Monday appears to have no bearing on whether it makes sense to kill whales rather than save them.

The above example is rather silly, but if you ask people to name our clearest instances of moral progress, they will inevitably point to changes that can be understood as cases of reduced arbitrariness: Why should whites enjoy freedoms that blacks do not? Why should men enjoy freedoms that women do not? Why should some tax-payers receive representation in government while others do not? And the same can be said for many of the moral battles currently being fought: Why should gays be barred from marriage and military service when heterosexuals are not? Why should people born in the United States enjoy its opportunities and advantages while those who were born elsewhere may not? Why should a relatively benign substance like marijuana be illegal while an extremely damaging substance like alcohol is not? And then there are questions that, at present, are primarily the concern of philosophers: Why is it acceptable to end the life of a fetus with poor life prospects, but not an infant with poor life prospects (Singer, 1995)? Why should one feel obligated to save a nearby drowning baby, but not a starving baby across the ocean (Singer, 1972)? Why should some people have greater prospects in life simply because they were born with greater talents or financial resources?[10]

---

[10] And just for good measure, there are cases of questionable arbitrariness that are fun to think about, even if they're hard to take seriously as practical issues: Why should individuals care about their own future selves more than those of other people?

Exploring these issues serves three functions. First, doing so may change our practices, as in the case of the civil rights movement. Second, in attempting to find a reasonable basis for what appeared to be an arbitrary distinction we may learn something about the facts and/or our values. For example, an inquiry concerning the treatment of animals may move us to learn more about the minds of animals. As a result we may learn something about animal psychology, our values, or both, even if we don't change the practice in question. Third, when our inquiry results in neither a change in our practice nor the identification of a satisfying justification for retaining it, we are humbled. As argued in the last chapter, our world suffers from moral overconfidence. We are likely to get along better if each of us recognizes that many of her moral practices and commitments are rather arbitrary.

Identifying and rooting out counterproductive and morally arbitrary behavior may seem like a rather narrow occupation for ethicists, and in a sense that is certainly the case. This kind of ethics is but one among many. However, from a practical standpoint, this project looms large. As noted above, most of the widely acknowledged moral progress we've made has come from asking questions such as these, and I see no reason why this trend should not continue.

5.1.4 Moral Psychology and the Taming of Moral Intuitions

Since Chapter 2 I have emphasized the relevance of moral psychology to our understanding of ethics. So far, psychology has been relevant in two principal ways: as an explanation for the apparent truth of moral realism (Chapter 3) and as a source of guidance in structuring our anti-realist practices (Chapter 4). Here, moral psychology makes a third contribution, as a partner in the project of applied ethics.

Normative ethics, in its most ambitious form, aims to organize all of our moral intuitions into a single, coherent moral theory. Once again, it fails because our intuitions are not sufficiently well-behaved, having been designed to serve our ancestors' reproductive needs rather than our present philosophical ones. Thus, instead of large-scale theory building, we must settle for something more modest, namely local consistency. But even within relatively circumscribed ethical domains, consistency is elusive, as illustrated above in the combined case of abortion, infanticide, and the killing of animals for food. In such cases one need not throw up one's hands, however. Instead, one may hope to disarm one or more of these unruly intuitions, either by altering them or finding reasons to discount them in one's decision-making.

Psychology is useful in this regard because psychology can tell us why we think the way we do, and knowing why one thinks something can change one's attitude toward that thought, or even stop one from thinking it. Take, for example, the incest case from Chapter 3 in which two adult siblings, Mark and

Julie, have sex using birth control. Most people say that their behavior is terribly wrong. Why? The evolutionary explanation is familiar enough. Matings between close relatives are especially likely to result in children with birth defects, making a powerful aversion to sex between close relatives an important biological advantage. Thus, we see incest as morally bad because it was, for our ancestors at least, biologically bad. Today we continue to see it as morally bad even in cases in which there is no biological disadvantage. Understanding this, what shall we say about Mark and Julie? On the one hand, we could stand by our intuitions and condemn them all the same. On the other hand, we could acknowledge that their behavior makes us a bit uncomfortable (if it still does), but conclude nevertheless that it's their business if that's what they want to do. The second response isn't *required*. Were we to condemn them for their incest, we wouldn't be making a *mistake*. Nevertheless, we're much less likely to condemn them once we understand of the etiology of our intuitive revulsion to their behavior.

Debunking intuitions through a better grasp of moral/evolutionary psychology will likely serve us well as we strive for $moral_2$ consistency by putting some distance between us and our intuitions. As noted above, we can make sense of the moderate liberal's respective intuitions concerning infanticide, abortion, and killing animals for food in evolutionary/psychological terms. As a result of understanding these intuitions, we may disown one or more of them. For example, we may conclude, as Peter Singer does (1979, 1996), that, in some

cases at least, it makes sense to end an infant's life or that it doesn't make sense to kill certain animals for food, and thus make our policies less arbitrary.

I should emphasize, however, that people are highly unlikely to turn to moral psychology in hopes that it will distance them from their intuitions. This is because most people are more committed to their moral intuitions than they are to making them consistent. People who think that allowing an infant to die is *absolutely positively wrong* are not especially interested in being talked out of that conviction, even if it would allow them to better defend their moral views. (We're talking about *values*, desires that people desire to retain (Frankfurt, 1971; Lewis, 1989).) The people whose moral commitments change as a result of their encounters with moral/evolutionary psychology are for the most part not people who turned to science in search of relief from inconsistent intuitions. Rather, they are people who encountered moral/evolutionary psychology for other reasons only to find that their understanding of moral/evolutionary psychology has undermined much of their common-sense moral thought. Thus, as a pedagogical matter, it seems that the way to get people to change or disown their moral intuitions is not to say, "Here's some science. It will change your moral outlook and make you, like me, have moral opinions that at present strike you as abhorrent." Rather, we friends of intuition-debunking through evolutionary/moral psychology are better off presenting the science as science and not as a device for turning common sense moralists into Peter Singer fans. (More on revisionist educational strategies later in Section 5.4)

The relevance of moral psychology to moral thought is ubiquitous, extending well beyond the cases I've mentioned. Consider our sense of justice. As noted above, our intuitions in the trolley and footbridge cases appear to depend on our emotional responses, which in turn may be understood in evolutionary terms. The violation one commits in the footbridge case is "up close and personal,";[11] one must push the innocent bystander to his death with one's bare hands. In the trolley case one simply hits a switch and lets the trolley do the rest. Why should we react differently to these sorts of behaviors? A hypothesis: "Up close and personal" violence was a serious issue for our hominid ancestors, while violence using elaborate contraptions that put fairly large distances between perpetrator and victim was not. It makes sense that we respond to these two dilemmas with different degrees of emotion, in spite of their similarity and in spite of the difficulties we face in explaining why we *ought* to treat them differently.

The trolley and footbridge cases are simple, but telling cases—fruit flies in the emerging science of evolutionary moral psychology. In these two cases one glimpses the broader organization of our moral thought and the various commitments we have as a result. As Singer (1972) noted decades ago, we who would without hesitation incur great risk and cost to ourselves to save a nearby drowning baby have no problem spending our money on needless luxuries when that money could save the lives of countless dying babies in other countries. Is it

---

[11] According to criteria outlined in Greene *et al*. (2001).

that we happen to care a great deal about physical proximity?[12] Might this not

have something to do with what sorts of things press our moral-emotional

buttons? Recall once again Josef Stalin's words: "A single death is a tragedy. A

million deaths is a statistic." How much death and destruction and plain-old bad

policy (Baron, 1998) has humanity suffered because our minds were not well-

adapted to preventing it?

The better we understand our moral intuitions, the more we are in charge

of them rather than the other way around. Our aim is not to destroy our

intuitions—What would we be without them?—but rather to *tame* them, to

understand them and accept them for what they are: not messengers of moral

truth, but biologically and culturally encoded dispositions that are at once the

basis for our deepest values and the source of much unnecessary suffering. Like

Neurath's (1959) sailors who must repair their boat at sea, we cannot step

outside of our values in our attempts to repair them. We cannot reject a portion

of our values without appealing to other values we hold. To what values do we

appeal at the highest level? Science cannot tell us the answer, but it can help us

understand the nature of the conflict within our values and in so doing direct our

evaluative judgment. I believe that scientifically informed moral thought does

have a normative direction, and I will say more about that in Section 5.3.

---

[12] Or is it the fact that these children are citizens of other countries? Not likely.

Would you, while vacationing in Paris, let a French baby drown in the Seine? For

a deeper exploration of these issues see Unger (1996).

5.2 Revisionist Moral Discourse

Now that we have a sense of what revisionist moral philosophers will be doing, what about the world outside of philosophy? After all, the revisionist aim is to get rid of realist moral thinking, not just in the halls of academia, but in the world at large.   To get rid of realist thought and the behavior it engenders, we can start by getting rid of realist talk.

In the previous chapter I sketched a plan for revising moral discourse: Speak of facts when possible, and speak of values in transparently subjective terms. Speaking this way would, I believe, help us resist the illusion of realism and thus further communication and understanding across moral lines. Earlier I said that some realist talk is worse than other realist talk when it comes to promoting misunderstanding and ill-will. In this section I explain why this should be so and which bits of language are most likely to be counterproductive.

5.2.1 Begging the Question with Deontological Language

Bernard Williams writes:

> Mackie's theory, and any like it, leaves a real problem of what should
> happen when we know it to be true... [I speculate] that the first victim of

this knowledge is likely to be the Kantian sense of presented duty… it is

the starkest example of objectification, and since there is nothing to it

except the sense of being given, it stands to suffer the most if that sense

is questioned. There are other ethical desires that are better adapted to

being seen for what they are. It is an important task for moral philosophy

to consider what they may be, and into what coherent pictures of ethical

life, philosophical, psychological and social, they will fit (1985).

Realist words like "good," "bad," "better," and "worse," while not ideal from a

revisionist perspective, are not so terrible either. First, this is because their

ordinary uses are more often subjective or hypothetical as compared to words

like "right" and "wrong." When I say that the sea bass was "good" last night, I'm

very likely expressing nothing more than that I enjoyed eating it, and when I say

that the movie last night was "better" than the one I saw last week, I may not be

expressing anything more than a preference. I *might* intend to express more

than that, but I might not. In contrast, if I say, as Bertrand Russell (1946) once

did, that bullfighting is *wrong*, I, like Russell, almost certainly intend to express

more than my distaste for it.[13] Second, our use of terms like "good" and "bad," as

---

[13] It's no accident that R. M. Hare's (1952, pg. 151) expressivist theory of moral

language begins with an analysis of the term "good" rather than the term "right":

"I have so far, in discussing the words used in moral discourse, confined myself

largely to the word 'good,' because the characteristics to which I wished to draw

attention are most easily illustrated in the behaviour of that word."

compared to "right" and "wrong," is more easily interpreted as making implicit reference to optional standards. When I say that ice cream is "good," I surely mean that it is good *for eating* (or, perhaps, *for tasting*) and not good for, say, cleaning my rug. In other words, our use of "good" in such contexts is implicitly *hypothetical*: *If* you want something tasty to eat, then ice cream is good. While terms like "right" and "wrong" and "rights" have some hypothetical uses of their own, as in the "right" way to hit a forehand volley, their uses are very often implicitly *categorical*. For example, when one says that it's *wrong* to kill people, it's not at all clear what, if anything, this is supposed to be wrong *for*. Wrong for *morality*? Yes, but standards imposed by morality aren't supposed to be optional. Indeed, we say that killing people is *just plain wrong*, regardless of one's goals or preferences. In contrast to our most moralistic uses of "right" and "wrong," it's harder to find uses of "good," "bad," "better," or "worse," that are so naturally interpreted as implying categorical demands. And without categorical demands, with only optional moral standards, there is no realism.[14] Thus, it seems that some bits of realist language are less subjective and more categorical than others, making them more realist and therefore more important to avoid.

The main enemy here, the language that is the most unambiguously categorical and objective in its implications, is the language characteristic of *deontology*. Deontology is usually contrasted with consequentialism. According to consequentialism, The Good precedes The Right. In other words, the moral

---

[14] *Pace* Harman (1996).

quality of an action is assessed in terms of its effects—or its intended effects, or its likely effects, or its perceived likely effects, or some other more sophisticated variant of this idea. Deontologists, in contrast, believe that The Right precedes The Good, or is at least somewhat independent of it. The moral quality of an action is not simply a function of its associated outcomes, but depends, rather, on whether that action exhibits proper respect for the dignity, freedom, autonomy, etc. of others in a way that goes beyond merely factoring in other people's well-being as part of the cost-benefit analysis. For consequentialists, every practical moral issue is largely an empirical issue, a matter of figuring out what states of affairs are likely to result from the action in question. For deontologists, many practical moral issues are to be settled *a priori* simply by considering the "intrinsic" nature of the action in question. In deciding whether it would be wrong to tell a certain lie, for example, a consequentialist's deliberation will consist of a tabulation of that lie's expected effects, both direct and indirect. In contrast, a deontologist may pass judgment on that action without much thought for its consequences. One might conclude, for example, that telling this lie would be wrong regardless of the good consequences it may achieve because all deception exhibits insufficient respect for the autonomy of others and is therefore wrong. Consequentialists and deontologists each help themselves to the full array of the common sense moral terms; "right," "wrong," "good," "bad," "ought;" but they emphasize them differently.

Another way to characterize the difference in emphasis is to say that consequentialism emphasizes the language of *balance* while deontology

emphasizes the language of *line drawing*. Line-drawing is a bad habit,[15] the normative expression of what is most destructive in moral realism. Recall from Chapter 3 the psychological basis of moral realism. People project their values onto the world, mistakenly thinking that the forceful intuitions they experience are perceptions of a mind-independent moral reality. How will people in the grip of such projections be inclined to express their views? Answer: with a language as forceful and absolute as their moral intuitions. Thus, an ardent liberal, one whose liberal judgments stem from deeply felt moral intuitions, will say that capital punishment is a clear violation of human *rights* and that abortion is a clear violation of a woman's *right* to choose. Likewise, an ardent conservative will say that family members of murder victims have a *right* to see justice served in the form of an execution and that abortion is a clear violation of a fetus' *right* to life. Of course, both sides will appeal to consequences when it serves them. Liberals will argue that capital punishment is not an effective deterrent and that making abortion illegal will only result in more backroom butchery and unwanted children. Likewise, conservatives will defend capital punishment's effectiveness as a deterrent and argue that illegal abortions and unwanted children can be replaced by adoptions. But the more honest members of either camp will admit that their commitments are not beholden to these sorts of contingent claims about the consequences of capital punishment and abortion. Once again, liberals—the

---

[15] Actually, it's not line-drawing *per se* that's bad. After all, there are many times one "has to draw the line." What is bad is appealing to drawn lines in attempting to settle moral issues. See below.

more ardent ones, anyway—oppose capital punishment because they *feel* that it is *just plain wrong*, likewise for conservatives and their opposition to abortion. This is why talk of *rights* is such a natural expression for these powerful commitments. It captures the clarity and absoluteness of the conviction, its insensitivity to the empirical calculus of costs and benefits. In other words, the notion of *rights* epitomizes the practical failure of moral realism, the stubbornness, rigidity, and irreconcilable differences that emerge when people believe that they, unlike their perverse opponents, *clearly* have the moral truth on their side.

There are two kinds of rights to which one can appeal, legal rights and moral rights, and in either case the appeal is inevitably question-begging. Gun-lovers appeal to their (legal) right to bear arms: "Read the Second Amendment!" But, of course, this gets them nowhere with their opponents, to whom the whole issue is whether we should change the Second Amendment or (re)interpret it so as to achieve the same thing. When two legal rights conflict, it's equally pointless to appeal to one of them as if this might settle the issue. This is especially true when there is no clear legal standard for determining which right takes precedence, but even when there is such a standard, one can question that standard on moral grounds just as one can question any law on moral grounds. The only time it's not question-begging to appeal to a legal right is when there is genuine ignorance of the letter of the law, as when board game players are forced, after much heated dispute, to consult the fine print on the inside of the box.

Given that moral realism is false, it is equally pointless to appeal to moral rights in the context of a purely moral disagreement. Suppose some people are discussing whether their mutual friend should confess to her husband that she's had an extramarital affair. "She *must* tell him. Gustave has a *right* to know!" one of them asserts. If realism were true, such a claim might convey useful information about the Moral Truth, but to people who know the terrible, horrible truth about morality, such assertions are no more effective than an expressivist "Boo!" or "Hurrah!," and probably less so given that revisionists will inevitably find people who speak as if they have the Moral Truth on their side somewhat annoying. Assertions about rights, wrongs, duties, and obligations (over)state a position, but they do nothing to *defend* it. They tell you what to do, but they don't tell you *why*, and they certainly don't give someone who is inclined to disagree a reason to change her mind. Contrast assertions about rights and other deontological entities with these contributions: "I think it's selfish to tell him. Let her live with it! Why make things worse for him?" This statement suggests a shift in perspective. Rather than conceiving of her confession as an act of bravery performed out of respect for her husband, it is portrayed as a selfish act, one of sacrificing her husband's happiness in order to ease her own sense of guilt. To this one might reply, "She should tell him because he probably already suspects, and knowing is better than always wondering." This comment performs a useful service by drawing one's attention to a relevant "empirical" issue, namely the possibility that her husband already suspects her infidelity and that he may have less to lose in being told about it than one might suspect. Another friend adds,

"Would you want to live a lie? I'd want to know the truth no matter what." This comment suggests yet another shift in perspective, one achieved through the empathetic exercise of role-reversal.   Unlike empty assertions about rights and duties, comments such as these do offer something of value. Some draw attention to relevant facts ("Good point! I hadn't thought of that!"), while others offer new perspectives on the same facts, inducements to view the same situation in a different evaluative light ("Good point! I hadn't thought of it *that way*!").

Moral realists, and those anti-realists who would emulate them, have the option of dogmatism, of blindly acting by moral norms that one takes to be authoritative. Revisionists, in contrast, have no choice but to acknowledge that all moral judgment is an imprecise process of weighing values. The nature of moral action requires the drawing of lines: One either jumps in and saves the drowning child, or one does not. One either votes to allow abortion or one does not. Of course, one will sometimes make compromises by adopting middle-of-the-road courses of action, but, at some level, all action is discrete. To any particular course of action one must say either "yes" or "no." Thus, while the inputs to moral judgment are fuzzy, fluid, and continuous *considerations*, the practical outputs of moral judgment are discrete *actions*. Deontology is intuitively appealing because it offers answers as clear and forceful as our intuitions, drawing *theoretical lines* that translate into *practical lines*, the kinds of lines that we, like it or not, are forced to draw by the nature of action. But, contrary to appearances, nature contains no true moral lines. We begin with only a mush of

morally relevant considerations, things we care about, and any lines that get drawn must be drawn by us. Therefore, any attempt to settle a moral question with deontological appeals to rights, obligations, etc. always begs the question. Such appeals are merely attempts to settle moral issues by insisting that they have, in effect, already been settled by Mother Moral Nature and the lines she has drawn.

In contrast to deontological language, characteristically consequentialist words like "good," "bad," "better," and "worse" may be used by revisionists without too much cringing because such words lend themselves to talk of *balancing*, to making sense of the moral mush rather than denying its mushiness. Justifying one's opposition to the action under consideration by labeling it "bad" is as bad as doing so by labeling it "wrong," but less dogmatic uses of "bad," "good," etc. can be useful. One can use such terms in one's subjective accounting of costs and benefits. When one says, for example, "It's *bad* enough that she did it... Why make things *worse* for him?" the "bad" and "worse" are not controversial labels that beg the ultimate moral question, but rather indications of uncontroversial[16] moral$_2$ sub-judgments that will feed into the all-things-considered judgment that practice requires.

Often one can think of words like "good" and "bad" as non-specific placeholders for evaluatively relevant factual terms. Continuing with the example

---

[16] When such judgments turn out to be controversial, then one must drop the evaluative label and go a level deeper, treating the use of the label as a moral issue in itself.

above, a revisionist might just as easily say, "Her actions have caused a lot of unhappiness already… Why create additional unhappiness for him?" Zealous revisionists may prefer a formulation of the latter sort, one completely purged of realist moral language, but the advantages here are minimal. Likewise, when one says that it's "better" to know than to wonder, one can easily interpret this as an empirical assertion, probably one to the effect that knowing the unpleasant truth makes one happier in the long run.

What about the "should" in "She *should* tell him, because…?" Here we might say instead, "I recommend telling him, because…" but, again, the gain is minimal. In this context the emphasis is on the factual claim following the "because," and not on abstract moral considerations.

What goes for private debates about marital infidelity goes for public moral debates as well. In the context of an openly anti-realist dialogue, what would it mean to say that a fetus has a *right* to life or that a woman has a *right* to choose? If all one means in saying these things is that one is against abortion, or in favor of allowing it, then why not just *say that*?[17] Packaging one's opinion as a claim about "rights" is just pointless propaganda. Perhaps, one might argue, that an appeal to a right can be understood as an appeal to a *default assumption*. To appeal to the moral right to free speech, for example, might be to appeal to the generally accepted principle that people should be able to say what they want in almost all cases. The problem is that in any real controversy in which "rights" are

---

[17] And if, for some strange reason, one wishes to *express* one's opposition to abortion without making a claim about oneself, then one can just say "Boo!" to it.

invoked, the question is inevitably about the *limits* of those rights. Therefore, it is pointless[18] for civil libertarians to defend flag-burning by appeal to the right to free speech, regardless of how natural this feels. Everyone is *generally* in favor of free speech. The debate is about whether to make an *exception* for this sort of speech. Pointing out that this case would be an exception does nothing to change the minds of those who want it to be an exception. In this case, as in others, appeals to rights are, once again, just question-begging propaganda, useless in the face of anti-realists who know the meta-ethical truth and aren't willing to play along.

In contrast, one can appeal to "good" and "bad" consequences without propagandizing. A pro-life conservative might say, for example, that restricting abortion rights is "good" (or "good" in certain respects) because it teaches people important lessons in personal responsibility. This claim is *not* empty. A young woman who gets pregnant as a result of irresponsible choices and is forced to carry a baby to term and give it up for adoption is likely to learn some valuable life lessons that she might not learn if given the option to end her pregnancy with a pill. This sort of claim contributes to a meaningful dialogue because it's something a pro-choice liberal can acknowledge. He can agree that there *is* an up side to anti-choice legislation, but then go on to argue that these advantages are outweighed by the disadvantages, "bad" things that an honest conservative can acknowledge as bad, such as harm caused by illegal abortions. Of course,

---

[18] Pointless so far as real dialogue is concerned. It may be useful as propaganda.

this kind of dialogue may not resolve the issue. There may remain untestable factual assumptions on either side concerning, for example, the existence of God and the nature of His will. Likewise, there may remain brute evaluative differences, for example, over the various weights to attach to the mutually acknowledged "good" and "bad" consequences of (dis)allowing abortion. But getting rid of question-begging talk of "rights" and establishing some common ground about advantages and disadvantages may help focus the issue. If, for example, the pro-choice camp could get the pro-life camp to acknowledge that its opposition to abortion is ultimately grounded in untestable religious beliefs *and nothing else*, that would be a very worthwhile achievement.

One might claim that I haven't given deontological language a fair shake. One might say that I've only considered the least sophisticated and most dogmatic forms of deontological language while ignoring the most dogmatic and simplistic uses of consequentialist language. To address the second charge first, my claim is not that all consequentialist moral language is good, or even that any of it is good. My claim is that *some* of it is *tolerable*, and therefore I'm rejecting, not ignoring, the bad uses of consequentialist language. The first charge is more serious. If what I'm really against is simplistic line-drawing rather than sophisticated balancing, then why is talk of balancing competing "rights" and "obligations" any worse than the equivalent consequentialist talk of balancing harms and benefits ("goods" and "bads")?

Some people who talk of balancing rights may think there is an algorithm for deciding which rights take priority over which. If that's what we mean by

"balancing rights," then we are wise to shun this sort of talk. Attempting to solve moral problems using a complex deontological algorithm is dogmatism at its most esoteric, but dogmatism all the same. However, it's likely that when some people talk about "balancing competing rights and obligations" they are already thinking like consequentialists in spite of their use of deontological language. Once again, what deontological language does best is express the thoughts of people struck by strong, emotional moral intuitions: "It doesn't matter that you can save five people by pushing him to his death. To do this would be a violation of his *rights*!"[19] That is why angry protesters say things like, "Animals Have Rights, Too!" rather than, "Animal Testing: The Harms Outweigh the Benefits!" Once again, rights talk captures the apparent clarity of the issue and absoluteness of the answer. But sometimes rights talk persists long after the sense of clarity and absoluteness has faded. One thinks, for example, of the thousands of children whose lives are saved by drugs that were tested on animals and the "rights" of those children. One finds oneself balancing the "rights" on both sides by asking how many rabbit lives one is willing to sacrifice in order to save one human life, and so on, and at the end of the day one's underlying thought is as thoroughly consequentialist as can be, despite the deontological gloss. And what's wrong with that? Nothing, except for the fact that the deontological gloss adds nothing and furthers the myth that there really are "rights," etc. Best to drop it. When deontological talk gets sophisticated, the thought it represents is either dogmatic in an esoteric sort of way or covertly consequentialist.

---

[19] See "The Trolley Problem" in Thomson (1986).

What about "ought" and "should?" These words are more or less equally comfortable in the mouths of deontologists and consequentialists, and, as we would expect, not so bad as "rights" and "duties," but a bit more dangerous than "good" and "bad." In asserted contexts one can avoid them by saying "I favor…," "I'm for," "I recommend…," "I support…," "I'm opposed to…," or "I'm against…" instead of "We ought to…," or "He should….," or "They must…." It's a bit hard to let go of these words in non-asserted contexts. For example how would a revisionist ask, "Should abortion be legal?" One could ask, "Are we for or against abortion?" but this sounds like a question about our current position, as if our votes have been cast and it's simply a matter of tallying them. "Shall we allow abortion?" and "Are we to allow abortion?" both sound a bit too realist, not to mention stilted. "Will we allow abortion?" makes it sound like the issue is one of predicting the future. As far as I can see, there is no standard English way of asking the anti-realist question, "Should$_2$ abortion be legal?" or of translating the phrase, "the dispute over whether abortion should$_2$ be legal." For this reason, revisionists may have to settle for "ought" and "should" for use in non-asserted contexts.

5.2.2 Calling it Like We See It: "True" Values and Morally Loaded Words

Ask a conservative why she supports tougher sentencing standards for juvenile criminals or opposes expanding welfare and she may say something about

what's "right" or "wrong," but more likely she will say something like this: "I support tougher sentences for juveniles and scaling back welfare because I *believe in responsibility*." Likewise if you ask a liberal why he supports lighter sentences for juveniles and the expansion of welfare programs, he will likely say that he *cares about people.* Liberals, of course, will *say* that they believe in responsibility, but, claim conservatives, they don't *really* believe in responsibility. Or, to put the point more clearly, what they believe in *isn't really* responsibility. They don't know what responsibility *really is.* Likewise, liberals will say that conservatives don't *really care* about helping people, in spite of what they say.

Conservatives know that liberals do not say, "Yep, that's us, the ones who don't believe in responsibility." Likewise, liberals know that conservatives don't say, "Yep, that's us, the ones who don't care about people." As a result, conservatives who loudly proclaim that liberals don't believe in responsibility may score points with other conservatives, but they do nothing to help liberals appreciate their conservative point of view. Liberals dismiss such accusations as ignorant nonsense, just as conservatives do when liberals claim that conservatives don't really care about people. Why do so many people, even those who would like to have a constructive dialogue, talk this way when it's so manifestly futile? As ever, the answer lies in moral psychology. Conservatives and liberals see the world in different moral colors.[20] A conservative looks at a juvenile criminal on the witness stand and sees, above all else, someone in

---

[20] Much of what I say in this section is drawn from George Lakoff, *Moral Politics* (1996).

desperate need of serious discipline—time for *tough* love. A liberal looks at the

same young man and sees, above all else, a product of unfortunate

circumstance, a bundle of anger and frustration crying out for attention—time for

tough *love*. To the conservative, it is perfectly clear that *responsibility demands*

tough sentencing, and when liberals wants to "coddle" young criminals instead of

sending them up for some hard time, the conservative concludes that liberals

must not believe in responsibility.   When responsibility stares them in the face,

they ignore it without apology. Nor do they believe in *true* compassion. If you

really care about the child, you don't *spare the rod*. At the same time, the liberal

wonders how anyone with even a *shred of compassion* could send this screwed

up kid who's trying to get by the only way he knows how to an adult prison where

he will be brutalized day after day, year after year. The liberal concludes that

conservatives *must not really care at all*. When hardship stares them in the face,

they ignore it without apology. And they don't really give a hoot about

responsibility either, the liberal says. If they did, they would acknowledge the

responsibility we all have to help our society's most troubled youth.

   Thus when conservatives say, "I'm a conservative because I believe in

responsibility," implying that liberals don't, and when liberals say, "I'm a liberal

because I care about people," implying that conservatives don't, both of them are

calling it like they see it. They don't realize that the values they are projecting

onto, say, juvenile criminals, the conservative's *to-be-punished-ness* and the

liberal's *to-be-cared-for-ness*, are mere projections. They don't realize that what

they see isn't really there. Liberals and conservatives alike see themselves as

doing nothing more than *heeding the call*, and those who do otherwise are either

blind to it or *willfully ignoring it*, as opposed to heeding some *other call.*

Conservatives think that if they just say the word "responsibility" enough times

they will get through, get those befuddled liberals to see what's staring them in

the face, namely *responsibility*. Liberals do the same thing with their favored

phrases and laugh when George W. Bush calls himself a "compassionate

conservative."

What to do about this? Answer: Understand what's going on and stop

talking past one another. Understand that moral realism is false and that there is

no *true version* of responsibility, compassion, freedom, or courage, etc.

Understand that in the context of political disagreement, all the most energizing

words, the one's that best express the intuitions one feels, mean different things

to different people and that in using these words one is using *mere words*. And

what holds for abstract words like "responsibility" and "compassion" also holds for

more concrete, but morally loaded words like "terrorist." Calling Osama Bin

Laden a "terrorist" is great for firing up Americans and generating domestic

support for a military campaign, but it does nothing to foster mutual

understanding between Americans and the *millions* of people who sympathize

with Osama Bin Laden. Make no mistake, I am not defending Bin Laden's

actions or adopting a "relativist" $moral_2$ position. I'm against Bin Laden, just like

you, Western reader. My point is simply that *insofar* as we are interested in

communicating with the multitudes who think Bin Laden has a point, we would be

wise to stop calling him a "terrorist," in spite of how perfectly justified it feels. Nor

does it make much sense to describe the Taliban as people who "hate freedom." Granted, their ideal society is objectively more restrictive than ours, but that doesn't mean that they "hate freedom." Every society restricts the freedoms it deems less important in order to maximize the freedoms it deems more important. The same reasoning that leads people to label the Taliban "freedom-haters" could be used by anarchists to say the same of us. Contrary to our rhetoric, we are not to be distinguished from the Taliban by our belief in freedom. They have their version, we have ours. We prefer ours. Indeed, we *deeply value* our kind of freedom. But nothing is gained when we claim proprietary ownership of the word.

Talking about what one "believes in" when talking about one's values is a bad idea in general because it suggests realism. Beliefs, unlike desires, are capable of being true or false, and therefore, in calling moral commitments "beliefs" one implies that one's own moral commitments are *true.* In general, I recommend replacing realist moral talk with talk about what one "cares about" etc., but, as I've argued, there are dangers here as well. One gets nowhere by saying or implying that the opposition doesn't care about things like responsibility, compassion, and freedom, even if that's what appearances suggest. Remember that what we see in our moral experiences are mere projections and that when we "call it like we see it" we end up talking past our opponents, pointing at things they just don't see, making them think that we're crazy and vice versa. Real political dialogue can occur only when people have a common currency for the exchange of moral ideas, when they can talk about

things they *both* can see. We turn to the establishment of this common currency in Section 5.3.

5.2.3 Show, Don't Tell: A Lesson from the Art of Fiction

Writers of fiction must shape their readers' opinions about the characters and events they describe. Beginning writers inevitably attempt to do this by *telling* their readers what to think, by describing characters and events using overtly evaluative language. But seasoned writers know that persuasion lies not in telling readers what to think, but rather in *showing* them what to think about: "Show, Don't Tell." "Don't judge." Simply describe the fictional facts as clearly and vividly as possible, and let the reader (gently guided by the author, always in the background) draw her own conclusions. Good writers know that moral persuasion does not require moral language, or even overtly evaluative language. Don't embellish. Don't editorialize. Just stick to the facts. Simple descriptions speak more eloquently than do torrents of praise and blame.

This makes sense given the psychology of moral development. We don't learn how to be moral$_2$ by being *told*, but rather by being *shown*. Adults who expect children to "Do as I say, not as I do," are bound for disappointment, and this should come as no surprise given the biological origins of our moral/social dispositions. Our primate ancestors, like chimpanzees, had only their actions and emotional expressions to guide their young. Moralistic language was not an

option. Like them, we learn from experience, and what a good novel or film does is give us *vicarious* experience, the more concrete and detailed the better. Moralistic editorializing is nothing more than a distraction.

There is an important exception to the "Show, Don't Tell" rule, and that is in the case of children's literature. According to the social intuitionist model, a child's moral sense develops largely in response to non-linguistic cues, but at the same time, we should expect children to be more susceptible than adults to the opinion-shaping influence of strong moral language.[21] Moral language persuades best when opinions are not yet formed, which is why writers of children's literature can get away with saying things like, "Mr. Billings was an awful, horrible man with a heart of stone." This sounds like a line from a children's book because it employs persuasive methods that, though appropriate for children, would insult the intelligence of most adult readers.[22]

---

[21] Adults respond to such language, too, but not in the same way. Their opinions are reinforced, not shaped, by strong moral language, except in cases concerning unfamiliar objects of interest, most notably individuals and groups. If I have no experience with Simon or with the Mormons, I'm likely to believe you if you say that they are "sleazy." If, on the other hand, you say that my friend Antoinette or high school teachers are "sleazy," I, having a firm, antecedent opinion to the contrary, I will think less of you, not them.

[22] Of course, adults can still appreciate this kind of writing, but to do so they must adopt a child's mindset. They have to "get into it."

Most moral discourse is the conversational equivalent of children's literature. Disputants speak to one another—or, rather, *at* one another—as if their interlocutors failed to pay adequate attention on the day elementary morality was explained. Unaware of the projective nature of value, they marvel at their opponents' blindness, their utter failure to see what is so perfectly obvious. Not knowing what else to do, they scold their opponents as if they were children, and scold them as if they were *belligerent* children when they fail to respond the first time.

What to do about this? Take a cue from good writers. Stick to the facts. Keep evaluative language to a minimum, and get rid of the most overtly judgmental, moralistic language. When it comes to changing the minds of people with well-entrenched opinions, the chisel is mightier than the sledgehammer.

Of course, might does not make "right."   If the persuasive methods employed by fiction writers are effective, it doesn't necessarily follow that we would be wise to promote their use in moral discourse. If what's lacking in ordinary moral thought is, among other things, a bit of healthy detachment from emotional moral intuitions, then the subtle, largely emotional persuasion of the fiction writer may be a step in the wrong direction.

I suspect that literary approaches to moral communication, like nearly all tools, may be used for good or ill. Using literary techniques to make the suffering of others vivid, for example, is likely to be a good thing. In contrast, the reinforcement of prejudices and knee-jerk moral responses through literary persuasion is unlikely to help. (For a sympathetic view of the literary approach to

public moral issues see Nussbaum (1995).) My point here, however, is simply

that the craft of fiction-writing suggests that effective moral communication does

not require, and may even be hindered by, overtly moral language.

5.2.4 Get Off Your High Horse: A Lesson from the Art of Negotiation and
Diplomacy

Based on the picture of moral psychology developed in Chapter 3, I have argued

that realist moral communication, though very natural, is likely to be

counterproductive. The effects of $moral_1$ thought and language on negotiation

and conflict resolution have not been studied scientifically, if only because one

does not ordinarily distinguish between $moral_1$ and $moral_2$ thought. A bit of

relevant empirical work has been done (Trainer, 1982), but most of what we have

to go on are the background considerations presented in Chapters 3 and 4 and

the commentary of experts in the fields of negotiation and diplomacy,[23] which,

---

[23] The experts in diplomacy and international affairs cited below are allied to

varying degrees with the "realist" tradition in political thought, not to be confused

with the tradition of "moral realism" in ethics. Their views have been much

criticized, but also, in my opinion as well as others', much misunderstood. I've no

intention at this time to defend their political worldview, the exposition of which is

often a regrettable muddle of description and prescription. I do, however, believe

that their hard-won insights concerning the destructive nature of moralistic

unfortunately, are based on personal experience rather than quantitative study. Nevertheless I believe their comments are instructive and suggestive.

From *Basic Negotiating Strategy* (1971) by Roger Fisher, director of the Harvard Negotiation Project:

Officials think of themselves as acting in morally legitimate ways. In order to get them to change their minds, we have to appeal to *their* sense of right and wrong. But this is the opposite of what most governments do. First they appeal to their own people's sense of right and wrong, attempting to whip up support by demonizing the opposition, and this may work. But then the opposition becomes much harder to deal with… We justify our actions in Vietnam by identifying our adversaries as evil aggressors acting in a grossly illegitimate way. They are murderers and violators of international law. They are behaving outrageously. Such characterization helps rally support for our side both domestically and, perhaps, among third parties. But it makes our adversary even less willing to listen to what we have to say. It makes it less likely that our adversary will make the decision we want (Pg. 112).

---

language and thought are valuable, though certainly not the last word on the matter.

An attempt to point out to an adversary that he ought to make a decision where the 'oughtness' is based on *our* ideas of fairness, history, principle, or morality is at best a diversion from the immediate task at hand; at worst it is destructive of the result we want (Pg. 113).

Governments too often regard their task as that of adopting a morally and politically justifiable posture or attitude rather than that of obtaining the best results that can be obtained under the circumstances (Pg. 17).

From *Beyond Machiavelli* (1994) by the same author:

A more useful question than "Who is right?" is: "Given these strong partisan perceptions, how can we move forward (Pg. 35)?"

In general, it is more useful to draft statements that describe feelings and the impact of what others do than to draft statements that judge or describe others (Pg. 28).

From Joseph Montville (1993) in *Conflict Resolution: Theory and Practice*:

Societies moralize to strengthen commitment to laws and codes and promote group cohesion. Countries with absolutist moralities makes resolution harder.

From "Assuming the Moral High Ground" in *Negotiation Across Cultures* (1991) by Raymond Cohen:

> American negotiators tend to be intensely irritated by Indian moralism, the assumption that India is somehow the repository of righteousness and objective truth in the world. Of course they are irritated: many Americans perceive that select role for themselves. Making this point, Paul Kreisberg, who served in India in both the 1950s and the 1970s comments wittily that negotiating with Indians is intensely frustrating because it is, 'like negotiating with yourself!'

From R. V. R. Chandrasekhara Rao (1976).

> When rivals contend in terms of moral attitudes not only does compromise become difficult but, more important, exposure of lapses leads to lack of credibility.[24]

From George F. Kennan in "Morality and Foreign Policy" (1985):

> Let us note that there are no internationally accepted standards of morality to which the U.S. government could appeal if it wished to act in the name

---

[24] Cited by Cohen (1991).

of moral principles. It is true that there are certain words and phrases sufficiently high-sounding the world over so that most governments, when asked to declare themselves for or against, will cheerfully subscribe to them, considering that such is their vagueness that the mere act of subscribing to them carries with it no danger of having one's freedom of action significantly impaired. To this category of pronouncements belong such documents as the Kellog-Briand Pact, the Atlantic Charter, the Yalta Declaration on Liberated Europe, and the prologues of innumerable other international agreements (Pg. 207).

[We should avoid] what might be called the histrionics of moralism at the expense of its substance. By that is meant the projection of attitudes, poses, and rhetoric that cause us to appear noble and altruistic in the mirror of our own vanity but lack substance when related to the realities of international life. It is a sad feature of the human predicament, in personal and in public life, that whenever one has the agreeable sensation of being impressively moral, one probably is not. What one does without self-consciousness or self-admiration, as a matter of duty or common decency, is apt to be closer to the real thing (Pg. 212).

From "The Necessity of Amorality in Foreign Affairs" (1971) by Arthur Schlesinger:

It is not only that moral principles are of limited use in the conduct of foreign affairs. It is also that the compulsion to see foreign policy in moral terms may have, with the noblest intentions, the most ghastly of consequences. The moralization of foreign affairs encourages, for example, a misunderstanding of the nature of foreign policy. Moralists tend to prefer symbolic to substantive politics. They tend to see foreign policy as a means not of influencing events but of registering virtuous attitudes (Pg. 73).

Laying down the moral law to sinning brethren from our seat of judgment no doubt pleases our own sense of moral rectitude. But it fosters dangerous misconceptions about the nature of foreign policy… A deeper trouble is inherent in the very process of pronouncing moral judgments on foreign policy. For the man who converts conflicts of interest and circumstance into conflicts of good and evil necessarily invests himself with moral superiority. Those who see foreign affairs as made up of questions of right and wrong begin by supposing they know better than other people what is right for them. The more passionately they believe they are right, the more likely they are to reject expediency and accommodation and seek the final victory of their principles. Little has been more pernicious in international politics than excessive righteousness (Pg. 73).

The primary motive [for getting involved in Vietnam], it seems probably in

retrospect, had little to do with national interest at all. It was rather a

precise consequence of the belief that moral principles should govern

decisions of policy. It was the insistence on seeing the civil war in

Vietnam as above all a moral issue that led us to construe political

questions in ethical terms, local questions in global terms, and relative

questions in absolute terms (Pg. 76).

The Indochina war was a morality trip, and moral absolutism was the final

stop. As early as 1965 the *New York Times* quoted an American pilot: 'I

do not like to hit a village. You know you are hitting women and children.

But you've got to decide that your cause is noble and that the work has to

be done.' In this anointed spirit we conceived ourselves the world's judge,

jury, and executioner and did our work in Indochina (Pg. 77).

An intelligent regard for one's own national interest joined to unremitting

respect for the interests of others seems more likely than the invocation of

moral absolutes to bring about greater restraint, justice, and peace among

nations (Pg. 77).

From Thompson's *Moralism and Morality in Politics and Diplomacy* (1985):

Another version of morality is that of Manicheim which portrays the world in radical terms of absolute good and evil or right and wrong. Americans are predisposed to a form of Manicheimism by childhood distinctions between good guys and bad guys or cops and robbers (Pg. 15).

Whatever the short-run advantages of [appeals to abstract principle], it has floundered in the long run because an individual or nation who claims an achieved morality that others have a duty to follow does so on an assumption of having itself attained moral perfection. Manicheimism is a false religion which sees the world as divided between good guys and bad guys and this disease has infected American thinking on foreign policy. Since World War I, we have divided the world into peace-loving and aggressor, freedom-loving and communist states and based foreign policy on such a distinction (Pg. 26).

5.3 Hurrah for Utilitarianism!

"Utilitarian: one who believes that the morally right action is the one with the best consequences, so far as the distribution of happiness is concerned; a creature generally believed to be endowed with the propensity to ignore their [sic] own

drowning children in order to push buttons which will cause mild sexual gratification in a warehouse full of rabbits"[25]

To a connoisseur of normative moral theories, nothing says "outmoded and ridiculous" quite like utilitarianism. This view is so widely reviled because it has something for everyone to hate. If you love honesty, you can hate utilitarianism for telling you to lie. If you think that life is sacred, you can hate utilitarianism for telling you to kill the dying, the sick, the unborn, and even the *new*born, and on top of that you can hate it for telling you in the same breath that you may not be allowed to eat meat (Singer, 1979).   If you think it reasonable to provide a nice life for yourself and your family, you can hate utilitarianism for telling you to give up nearly everything you've got to provide for total strangers (Singer, 1972; Unger, 1996), including your own life, should a peculiar monster with a taste for human flesh have a sufficiently strong desire to eat you (Nozick, 1974). If you hate doing awful things to people, you can hate utilitarianism for telling you to kidnap people and steal their organs (Thomson, 1986). If you see the attainment of a high quality of life for all of humanity as a reasonable goal, you can hate utilitarianism for suggesting that a world full people whose lives are barely worth living may be an even better goal (Parfit, 1984). If you love equality, you can hate utilitarianism for making the downtrodden worse off in order to make the well off even better off (Rawls, 1971). If it's important to you that your experiences be

---

[25] From Henry Fitzgerald, "A Non-Philosopher's Guide to Philosophical Terms." (http://philrsss.anu.edu.au/~henry/gloss.html as of 3/1/02)

genuine, you can hate utilitarianism for telling you that no matter how good your life is, you would be better off with your brain hooked up to a machine that gives you unnaturally pleasant artificial experiences. No matter what you value most, your values will eventually conflict with the utilitarian's principle of greatest good and, if he has his way, be crushed by it.[26] Utilitarianism is a philosophy that only… well, only a *utilitarian* could love.

With some trepidation, and like many who have come before me,[27] I will defend a utilitarian approach to morality, but in a non-standard way.[28] I do not claim that utilitarianism is *true* or *correct*. Nor do I claim that it is somehow built into or presupposed by our moral concepts or language (Hare, 1981). (In other words I reject both the realist-cognitivist and expressivist versions of utilitarianism.) Nor do I recommend a utilitarian approach to every domain of moral life. Instead, I modestly recommend utilitarianism as a *public standard*[29] for evaluating actions and policies, as a meta-policy that is likely to meet the

---

[26] Unless, of course, your values happen to be perfectly utilitarian.

[27] To name a few: Bentham (1982), Mill (1998), Sidgwick (1890), Smart (1961, 1973), Brandt (1979), Singer (1979), Hare (1981), Wright (1994).

[28] For my purposes there is no need to settle on a particular version of utilitarianism, but the type I have in mind is a hedonic version according to which the best action is the one that maximizes the overall level of happiness. I will use "happiness" and "utility" interchangeably. The defense of utilitarianism I offer is similar in spirit to Robert Wright's (1994).

[29] See also Goodin (1995).

needs of people who have come to terms with the metaphysical, psychological, sociological, and evolutionary truth about morality. Utilitarianism, I claim, follows not *logically* from the truth about morality, but *psychologically*. It's what you're likely to want once you know the truth.[30] This, of course, is an empirical claim, but given the dearth of psychologically and metaphysically informed moral thinkers to serve as subjects, my speculations will have to suffice for now.

In the last section I explained why deontological talk is likely to fail in a psychologically informed, anti-realist environment. Once we've put some distance between ourselves and our intuitions, once we know where our intuitions come from and understand that they are not messengers of moral truth, drawing deontological moral lines and expecting others whose inclinations are otherwise to respect those lines just seems silly. And with deontology and the realism it reflects discredited, the utilitarian approach is, I claim, simply what's left. Nearly everyone is a utilitarian to some extent. Nearly everyone agrees that *all other things being equal* raising someone's level of happiness, either your own or someone else's, is a good thing, and that lowering someone's happiness is a bad thing.[31] More specifically, utilitarian harms and benefits matter to everyone, and in the same general way, thus making them a badly needed common

---

[30] See Blackburn (1993a) on affinity between projectivism and consequentialism. In claiming this I am not, in a roundabout way, claiming that utilitarianism is true. See the final footnote of this section.

[31] Some people would say, for example, that increasing the happiness of someone who deserves to suffer for her crimes is not a good thing.

currency for moral exchange. As noted above, pro-choice liberals can acknowledge that restricting abortion rights would have *some* good effects, just as pro-life conservatives can acknowledge that keeping abortion legal would have *some* good effects. They may disagree about the balance of these harms and benefits and about what other considerations are important, but when they talk about harms and benefits in terms of human happiness and suffering, they are speaking *to* and not *past* each other.[32]

But in spite of the virtually universal acceptance of utilitarian currency, almost no one is a full-blown utilitarian. To earn that distinction one must not only take utilitarian considerations into account; one must take *nothing else* into account. That is, one must assess all actions and outcomes solely in terms of their effects on the total level of happiness, and doing this often leads to moral judgments that *seem wrong*, sometimes *dreadfully wrong*, as suggested above. Utilitarians typically defend their position against such common sense objections in two main ways. First (Strategy 1), they argue that often utilitarianism merely *seems* to have counterintuitive implications. Second (Strategy 2), utilitarians attempt to discredit anti-utilitarian moral intuitions by interpreting them as the

---

[32] Some people have worried about "the problem of interpersonal comparisons of utility," but this is really a problem of precision and certainty rather than a problem with the meaningfulness of such comparisons. Interpersonal comparisons of utility are perfectly commonplace, as when one decides to give a spare theater ticket to one friend rather than another because one believes she will enjoy it more than the other. See Harsanyi (1982).

products of reasonable moral dispositions or heuristics that have been

overextended.   I will employ a bit of each of these strategies, but I will also

introduce a new tactic (Strategy 3), one grounded in the meta-ethical conclusions

reached in Chapter 2.   If moral realism were true, we would presumably aim to

act in accordance with the true moral theory. Thus, under realism, to defend

utilitarianism is to defend it as *the truth*, and this means defending it against *all*

common sense intuitive objections, even ones concerning unrealistic hypothetical

cases. [33] But in rejecting realism two important things change. First, our

intuitions can no longer be seen as messengers of moral truth.[34] The extent to

which we take them as authoritative is up to us. Second, intuitions concerning

unrealistic cases need not be terribly important.[35] A practical guideline need not

---

[33] Likewise, if utilitarianism were somehow built into our moral language or

concepts, proper use of our language/concepts would require an all-

encompassing utilitarian outlook.

[34] Of course, one can, in principle, be a realist and think that one's moral

intuitions are not even reasonably good guides to moral truth, but I know of no

one who adopts this unusual view. And this makes sense given the

psychological account of moral realism offered in Chapter 3: We're realists

because we find our intuitions so compelling.

[35] Take, for example, a moral theory according to which an action is wrong if

everyone's performing that action would lead to significant harm. One might

object to this view by pointing out that it would lead to significant harm if

everyone flushed her toilet at the stroke of midnight, overburdening the sewage

be an iron-clad rule that we would want to live by in all possible worlds; it just

needs to do the job here and now, and therefore the standards for evaluating

practical moral guidelines are far less demanding than those for evaluating moral

theories that purport to be true.[36] Thus, once we've abandoned our realist

ambitions, we can dismiss objections to utilitarianism (and other theories) that

are not sufficiently realistic. I will argue that utilitarianism, unlike other theories,

survives the transition from candidate moral truth to candidate practical guideline

very well.

The first main objection to utilitarianism is that it would have us do

intuitively awful and unfair things. Take, for example, Thomson's (1986)

transplant case. You're a doctor with five patients, each of whom is about to die

due to some kind of organ failure. In walks a healthy person whose organs you

---

system and causing dangerous floods, even though our moral sense tells us that

flushing one's toilet at the stroke of midnight is not wrong. *If* we were in search of

the moral truth, we'd have to find a new theory (or modified version of this one)

that circumvents this gimmicky counterexample. However, if we're just looking

for a *practical guideline*, then we can ignore this objection because it's

exceedingly unlikely to be a problem in the real world.

[36] Less demanding in some ways, more demanding in others. Indirect principles

such as "Do that of which we would approve if we were fully informed and fully

rational," do very well when it comes to avoiding troubling counterexamples

because they don't offer any practical guidance, but for the same reason they fail

miserably as practical guidelines.

could harvest against his will and then distribute to the other five people, saving

five lives at the expense of one. Most people have a clear intuition that it would

be wrong to do this, an apparent strike against utilitarianism. In response to this

objection, a utilitarian may employ Strategy 1. Utilitarianism doesn't really tell

you to perform this secret operation because it is so unlikely to achieve a net

benefit. If word of your clandestine activities were to get out, panic would set in

around the nation, if not the world. Hundreds, maybe thousands, of people

would die or suffer serious health problems for fear of visiting their doctors. And

how would you guard your secret? Would you perform five transplant operations

by yourself? And won't people wonder what happened to the guy who died, or

why his body is missing half his organs? And won't the organ recipients wonder

where their organs came from? And so on, and so forth. One might object that

all of these practical considerations are beside the point. The fact that you

wouldn't be able to get away with this operation isn't what makes it wrong, and

thus, the objection goes, utilitarianism gets the right answer, but for the wrong

reason. To put the objection another way, one can imagine a situation in which

these practical difficulties can be overcome. Suppose, for example, that Clark

Kent had gone into medicine instead of journalism. Intuitively, it would still be

wrong for him to perform this operation, but utilitarianism says otherwise, a clear

strike against it. Now we jump to Strategy 3. Forget about Superman. Forget

about doing the "right" thing for the "wrong" reason. We're just looking for a

practical guideline, not an eternal moral code that handles fantasy cases as well

as real ones or that tells us which among our reasons for action are

philosophically privileged. And, moreover, we're looking for a workable *public* standard.[37] Even if sometimes, under bizarre circumstances, it would better to perform this sort of secret operation, we would clearly be worse off if our laws gave health-care providers that sort of discretion.[38]

This utilitarian one-two punch works against similar objections. Suppose there is an epidemic of crime, and that the police could get a lot of mileage out of a high-profile conviction, an example to hold up to other would-be criminals and those who've lost faith in law enforcement. Suppose further that the police have in their custody a man who is almost certainly guilty, but not enough evidence to convict him. Why not manufacture some? He's probably guilty anyway, and the conviction would do a lot of good. Applying Strategy 1, we point out the practical difficulties in doing this and the awful risk one would take in trying. Surely the man will claim that he'd been framed. Surely some journalists will look into it. Surely the secret will come out eventually, eroding the public's trust even further, etc. Applying Strategy 3, we point out that the fact that utilitarianism would endorse such behavior under highly unrealistic conditions is irrelevant. Moreover, we point out that a publicly acknowledged standard allowing law enforcement officials to manufacture evidence at their own discretion would

---

[37] This constraint is essentially the same as Rawls' (1971) "publicity condition."

[38] This invocation of "rule utilitarianism" does not suffer from the usual problem of collapsing into "act-utilitarianism" because rules of the sort that we could actually use as public standards must be relatively simple and not allow officials too much discretion over their interpretation.

clearly *not* serve the greater good, thus making this charge irrelevant for the purposes of creating public standards for moral, legal, and political decision-making.

Take another famous objection to utilitarianism, Nozick's (1974) famous "utility monster" objection. Suppose there is a creature that gets incredible amounts of utility out of gobbling people up. According to utilitarianism, we would each be obliged to feed ourselves to this monster so long as the utility he gains from eating us is sufficiently large to outweigh the utility losses we incur. Once again, as anti-realists in search of a useful practical guideline rather than a true theory, we are free to say, "Who Cares? There aren't any utility monsters around here, nor are there ever likely to be."

I believe that this pattern is quite general. Our intuitions are not utilitarian, and as a result it is often possible to devise cases in which our intuitions conflict with utilitarianism. But at the same time, our intuitions are somewhat constrained by utilitarianism. This is because we care about utilitarian outcomes, and when a practice is terribly anti-utilitarian, there is, sooner or later, a voice in favor of abolishing it, even if the voice is not explicitly utilitarian. Take the case of drunk driving. Drinking is okay. Driving is okay. Doing both at the same time isn't such an obviously horrible thing to do, but we've learned the hard way that this intuitively innocuous, even fun, activity is tremendously damaging. And now, having moralized the issue with the help of organizations like Mothers Against Drunk Driving—what better moral authority than *Mom*?—we are prepared to impose very stiff penalties on people who aren't really "bad people," people with

no general anti-social tendencies. We punish drunk driving and related offenses in a way that appears (or once appeared) disproportionately harsh because we've paid the utilitarian costs of not doing so.[39] The same might be said of harsh penalties applied to wartime deserters and draft-dodgers. The disposition to avoid situations in which one must kill people and risk being killed is not such an awful disposition to have, morally speaking, and what could be a greater violation of your "rights" than your government's sending you, an innocent person, off to die against your will?[40] Nevertheless we are willing to punish people severely, as severely as we would punish violent criminals, for acting on that reasonable and humane disposition when the utilitarian stakes are sufficiently high.[41]

---

[39] A campaign against talking on the phone while driving is underway. See http://cartalk.cars.com/About/Drive-Now/.

[40] For any given individual it's only a *risk* of death, but it's still a severe imposition.

[41] Another example is our policy of non-negotiation with politically motivated kidnappers. As I write, an American journalist is being held hostage by a group of Islamic militants in Pakistan. They say they will kill him if their demands are not met, but the U.S. government has refused to negotiate with them. As much as we would like to save this man's life, we are willing to let him die in order to stand by our policy which, though emotionally difficult to live with, is bound to save more lives in the future. At least that's the idea. (Note: Journalist Daniel Pearl was eventually killed by his kidnappers.)

Drunk-driving and the shirking of military duty are real-world problems that produce serious harms, and therefore it's no surprise that we have strong prohibitions against them in spite of their being born of dispositions that are really not so bad. Drunk-driving is a relatively recent phenomenon, and therefore the prohibition against it rests on a somewhat thin, intellectual appreciation of a cause-effect connection. The prohibitions against desertion and draft-dodging rest on moral counter-dispositions that run much deeper culturally, and probably biologically.[42] The crucial point illustrated by both cases, however, is that big utilitarian considerations tend to work their way into our moral norms *provided that they concern real-world phenomena from which we can learn*, either through explicit learning or through cultural and/or biological adaptation, although this is certainly not always the case.[43] And this is why doing the utilitarian thing is less often counter-intuitive in real-world situations as compared to fantasy cases, cases concerning bizarre situations in which the cues that set off our moral alarm bells are curiously divorced from the consequences that usually accompany them.

That said, there are still plenty of real-world cases in which our intuitions are anti-utilitarian. Consider the proviso "provided that they concern real-world

---

[42] See Fehr and Gachter (2002).

[43] The clearest sort of exception to this principle are situations that resemble those in the prisoner's dilemma, situations in which the optimal strategy for achieving individual survival/happiness leads to outcomes that are far from optimal in terms of collective survival/happiness. See Section 5.4.2.

cases from which we can learn." Learning takes time, and we now face a great many relatively new problems. As I argued in Chapter 4, the most general new problem we face, the subject of this essay, is that of dealing with our moral differences now that we live in a world in which people with very different moral sensibilities must coexist. We can view this as a problem with the newness of the "we" who must learn. For most of human existence "we" was a village. Now, to use a once popular phrase, "We are the world," in all of our glorious and troubling diversity. Unlike personal violence and other classic moral violations, global warming and the proliferation of weapons of mass destruction etc. pose new problems, problems that may impose utilitarian costs like the world has never known. Fortunately, we (most of us, anyway) have yet to pay those costs. The flip side, however, is that we've yet to learn how to avoid these costs the way we usually learn, i.e. the hard way. Our evolving moral sensibilities haven't had time to absorb the utilitarian costs associated with these behaviors. (See Section 5.4.)

Above I gestured toward behaviors and policies that seem acceptable, but that really aren't, so far as long-run human happiness is concerned. Likewise, there are actions and policies that strike us as unacceptable, but that really aren't so bad from a utilitarian point of view. Once again, these tend to be in response to relatively new moral issues, cases in which technology has given us an option that's been unavailable for most of human history, an option that is very useful but morally repugnant or, in other cases, an option that feels morally obligatory in spite of its lack of utilitarian value. In dealing with such discrepancies between

common sense intuition and utilitarianism, I invoke Strategy 2, the psychological

debunking of intuitions. In section 5.1 I explained how a better understanding of

our moral intuitions and where they come from can change the way we think

about a wide range of issues from consensual incest to bioethical issues such as

abortion, vegetarianism, and infanticide to matters of personal harm of the sort

dramatized in the trolley and footbridge cases. I believe that people who

understand the history behind our moral intuitions are more likely to reach

utilitarian conclusions concerning such cases. Take Peter Singer's (1979, 1995)

suggestion that the parents of an infant with a severely debilitating and incurable

disease ought, in some cases at least, to be allowed to end their infant's life.

Many people find this suggestion appalling. They have strong moral intuitions

that tell them that ending an infant's life, even in circumstances such as these, is

just plain wrong. But those of us who understand the etiology of these intuitions

and know the meta-ethical truth will likely take a different view. We will

understand why natural selection, reinforced by cultural norms, would furnish us

with strong reactions against ending the lives of infants under all

circumstances.[44] We also know that there's no such thing as a *right* to life and

that life is not *sacred*. Having no true moral principles to guide us, we will simply

consider the things we care about. Do we care about preserving life? Not life

---

[44] Intuitions against killing one's own baby may not be as widespread as we

Westerners tend to think. Infanticide has been practiced in many cultures,

although it is less common in cultures in which resources are more abundant

such as our own.

per se. We' re willing to kill animals when the stakes (and in some cases

"steaks") are pretty low. Is it *human* life that matters? Knowing that humans are

not metaphysically privileged, this seems rather arbitrary. As anti-realists we

must presume that the universe has no special preference for *Homo sapiens*.

Perhaps it's some characteristic that humans happen to have that matters,

something that other species could have but don't. The problem is that those

sorts of characteristics (rationality, self-awareness, etc.) seem to be

characteristics that infants lack, especially ones with severe mental disabilities.

One thing we *do* care about is the happiness/desires of the relevant parties. If

caring for a severely disabled infant would be a great burden to its parents,

especially when they could devote their time, love, and resources to a healthier

child, then surely that is at least a point in favor of letting them move on with their

lives, if that's what they want. What about the infant? If the child's life is going to

be a drawn-out painful struggle culminating in an early death, then is it really in

that child's interest to live? And what about the interests of the child or children

that would live in its place? And when it's not clear what would actually promote

the greatest happiness, isn't it best, as a general rule, to leave the decisions up

to the people whose lives are most affected by them, so long as they are capable

of making them? Taking all of these considerations into account, it does not

follow that a revisionist anti-realist must, on pain of inconsistency, come to

utilitarian conclusions. She could embrace the intuition that ending the life of an

infant is always *just plain wrong*, even while knowing that this intuitive judgment

is not true. But, for the reasons I've given, I believe this is unlikely to happen.

There are other anti-utilitarian intuitions with which one must contend. Perhaps the most influential critique of utilitarianism is John Rawls' in *A Theory of Justice* (1971). Elsewhere I argue that Rawls' critique is based on a mistaken understanding of utility, a mistake that Jonathan Baron and I (2001) documented experimentally. Other influential critiques are Robert Nozick's (1974) "experience machine" argument and Derek Parfit's (1984) argument concerning the "repugnant conclusion." Elsewhere (Greene, 2001) I deal with these objections as well, primarily through psychological debunking of the intuitions on which they depend. There is, however, one important objection to utilitarianism that will not go away so easily. This is the objection that utilitarianism is *too demanding*. We middle class Westerners have, through our disposable income, the ability to greatly improve the lives of many of the world's less fortunate people, and since there are so many people who need our help, it follows that anyone who seriously intends to maximize utility will have to devote the bulk of her resources to charitable activities, forgoing not only modest middle class luxuries such as dining in restaurants, but more significant things as well, such as the pursuit of a career that is enjoyable but not maximally lucrative, devoting resources to one's family, or even having a family. Strictly speaking, utilitarianism demands that you turn yourself into a utility-maximizing machine, which in today's world probably means devoting all of your energy and resources to raising money for strangers.

Most people, even people who are generally sympathetic to utilitarianism, are not inclined to live like this. And this makes perfect evolutionary/psychological sense. We were not designed to be impartial

promoters of happiness. We were designed to spread our genes by serving

ourselves and, to a lesser extent, others, especially kin and those with whom we

are likely to have personal interaction. Altruism may have a strong innate basis

(Dawkins, 1976; Axelrod, 1984; Wright, 1994: Sober and Wilson, 1998), but we

are by no means designed to "love thy neighbor as thyself," especially our

"neighbors" on the other side of the world. But in this case, knowing the scientific

story behind our less-than-fully-utilitarian intuitions doesn't do much to debunk

them, at least not in practice. I know that I'm selfish, and I know why I'm selfish,

but that doesn't change the fact that I'm *still selfish*. I give some money to

charity, but I could certainly give more, and as I look to the future I know that I am

unlikely to sacrifice most of the things I want for myself, from eating sushi to

having a family. And yet my utilitarianism demands it. The same problem arises

in the public domain. How much of our national budget shall we devote to

foreign aid? Whatever answer utilitarianism demands, it is surely far more than

our nation would tolerate giving any time soon. Given its unreasonable

demands, how can we endorse utilitarianism, even as a practical guideline for

public policy?

In a sense we can't. We may have to make an *ad hoc* exception to our

utilitarian guidelines in order to account for the fact that, when it comes to the

flow of resources to people outside of our privileged circles, our values are

deeply un-utilitarian. But here's another way to look at it. Utilitarianism tells us to

increase the world's level of happiness *as much as possible*. We readily

acknowledge *external* constraints that narrow the range of possible actions; why

not *internal* ones? We wouldn't fault ourselves for not performing a certain action if something external such as the weather made it impossible. Perhaps we can view the constraints imposed upon us by our psychology in the same light. We can't be perfectly impartial in doling out our resources because it's simply not in our nature, but we can do our best to be as impartial as *humanly* possible, and to support institutions that foster greater impartiality. Thus, we may maintain utilitarianism as our ideal, understanding that we will never live up to it.

This separation of ideals from actions is not the cop-out it may appear to be. Maintaining an ideal from which one intends to deviate is more than saying, "It would be nice." Take the ideal of perfect honesty. Unlike Kant (1993), I don't want people to be so thoroughly honest that they would refuse to lie in order to save a friend's life, and for this reason I cannot honestly say that I maintain perfect honesty as an ideal. In contrast, if I could somehow rewire the human mind so that each of us would love one another as we currently love ourselves, I would do it. Thus, in this admittedly impractical sense, I can maintain the impartiality demanded by utilitarianism as an ideal, even if one has no intention of being perfectly impartial.

Note also that holding a principle as an ideal from which one intends to deviate is not a moral blank check. There is a difference between trying unsuccessfully to live up to an ideal and not even trying because one knows that one will not be able to so. When someone discards the request from Oxfam and says "I just can't bring myself to do it," it may be difficult to tell whether that person has made a sincere effort or simply used a convenient excuse, but that

doesn't mean that there's no fact of the matter. Being a good utilitarian is tough, and, as ever, one must be one's own conscience.

Is this a recipe for madness? Constantly wrestling with one's most natural human impulses? The problem is perhaps most stark when it comes to one's personal relationships. Shall I pay for my child's birthday party only after finding that I *just can't bring myself* to give the money to charity instead? It's best, I think, to know one's limitations and live relatively guilt free within them. Making one's life a series of painful moral battles is unlikely to produce the best results in the long run, anyway. One must strike a balance between struggle and livability in one's attempts to live the utilitarian life.

Utilitarianism is not true, nor is it an adequate systematization of anyone's values. No one is a utilitarian at heart. But at the same time, everyone is a utilitarian in part. As realists, relying on our moral intuitions as guides to moral truth, we have every reason to reject utilitarianism. But as people who know the truth about morality in search of practical moral guidelines that people with different moral outlooks can share, utilitarianism is, I think, our best hope.[45]

---

[45] I've argued that people who know the facts are likely to converge on utilitarianism, or at least on utilitarian values. Have I not undone my anti-realism? Am I not saying that utilitarianism is, after all, the *correct* view, that it is, in the end, *true*? No. Being true is not the same as being that upon which we would agree under favorable conditions. According to Smith, what's right is what we would desire if we were fully informed and fully rational. I'm happy to suppose that a world organized around utilitarian principles is what we would desire if we

5.4 The Tragedy of Common Sense Morality and the Scientific Worldview

"…but natural selection favors the forces of psychological denial. The individual benefits as an individual from his ability to deny the truth even though society as a whole, of which he is a part, suffers. Education can counteract the natural tendency to do the wrong thing, but the inexorable succession of generations requires that the basis for this knowledge be constantly refreshed."

—Garrett Hardin, "The Tragedy of the Commons" (1968)

I've argued that moral realism is perfectly natural, encouraged by the design of our minds. At the same time, I've argued that moral realism, at least in our

---

were fully informed and fully rational, but I deny that that this makes utilitarianism true for reasons given in Chapter 2. Similar remarks apply to Scanlon's (1982) indirect theory of what is "right." See also Scanlon (1998).

More importantly, I don't think we gain anything by calling utilitarianism "true" whether or not it is. As long as people's values are anti-utilitarian they will simply deny my empirical claim that knowing the facts makes one's values utilitarian, and once people know the facts they will, if I'm right, have utilitarian values, at which point it doesn't really matter whether or we call the values we share "true" or just "our values."

present context, is extremely pernicious, exacerbating conflict and promoting

miscommunication both within and across national borders. To remedy this

problem, I've suggested that we stop talking and thinking like realists. But how is

that going to happen? Getting a few philosophically inclined intellectuals to

abandon common sense moral thought and language is not an outrageous goal,

but what about people in general?

I think it's useful to analyze this problem in decision-theoretical terms. The

problem of moral realism is an instance of Hardin's "tragedy of the commons."

Tragedy ensues when individuals are free to overuse a public resource, as, to

take Hardin's example, when each of many herdspeople, all of whose livestock

graze on a common pasture, is free to increase the size of his flock. When a

herdsperson adds an animal to his flock, the benefits of raising and selling that

animal accrue solely to him, but because his animals graze on a common

pasture the costs of raising that animal are spread throughout the community.

Because each herdsperson receives the full benefit of having an additional

animal for himself but pays only a fraction of the cost, it is in each herdsperson's

interest to add to his herd. And since every herdsperson is in the same situation,

each pursuing his own self-interest will destroy their common resources, leaving

all of them with nothing. Such commons problems are all too common. For

example, in the case of pollution, the individual polluter receives the full benefit of

his action (getting rid of his pollutants), but pays only a fraction of the cost (the

harm to the environment). As a result, pollution is profitable for the individual

polluter, and, without some means of changing people's incentives, the world grows increasingly polluted.

Moral realism is, I claim, a heretofore unrecognized form of pollution. When individuals think they have the Moral Truth on their side they are stubborn and uncompromising, making it more difficult for them to reach mutually satisfactory agreements with those around them. At the same time, being a moral realist—or, at least, being perceived as one—may give one a strategic advantage.[46] Individuals who can convince others that they are unwilling to compromise, perhaps by *being* unwilling to compromise, are more likely to get what they want in a wide variety of strategic situations. For example, if you can convince your car salesman that you will settle for nothing less than your named price, and she can give you your named price, you're in very good shape. And aside from its potential strategic advantages, moral realism *feels good*. It pleases us to believe that we are morally right and that the opposition is morally wrong. Getting one's *just deserts* is more satisfying than simply getting what one wants, and, even more to the point, giving the opposition their *just deserts* is more satisfying than simply preventing them from getting what they want. (A recent study strongly suggest that people punish altruistically, that is, punish others even when punishment carries a known net cost to the punisher (Fehr and Gachter, 2002).)

Thus, realist individuals pollute the world with stubbornness and self-righteousness or, if you prefer, deplete the world's supply of reasonability and

---

[46] See Schelling (1960).

willingness to compromise. What to do? Commons problems may be solved by one of two means, political and moral. Political solutions are reached when a governing body changes people's incentive structures by punishing overuse of common resources. Clearly, a political solution to the tragedy of common sense morality is not available. We cannot make realist thought and language *illegal*. That leaves only moral solutions, which are achieved when individuals decide to change their behavior on their own, presumably out of a sense of moral obligation to the community. When it comes to large-scale commons problems, political solutions are generally more effective than moral solutions, but moral solutions do work from time to time. Could there be a moral solution to the problem of moral realism? Paradoxical as this may sounds, I think the answer is yes.

And we need not start from scratch. Many people resent *moralizing*. We are often resentful when people attempt to impose their moral views on others. (See Section 5.2.4.) Of course, we all do this to some extent, if only implicitly and tacitly. Every law is a restriction of freedom, and therefore anyone who supports the rule of law imposes her moral views (*moral$_2$* views, at least) on other people. But some people, we think, go too far. They take their own morality a little too seriously by expecting others to take it as seriously as they do.

We can, I believe, build on this resentment of moralizing. The new twist is the realization that *one person's morality is another person's moralizing.* When liberals encounter conservatives who want to put prayer back in public schools or oust a sitting president for lying about his sex life, they see this as bald faced

moralistic meddling. Likewise, when conservatives encounter liberals who want take John Q. Tax-Payer's *hard-earned money* and use it to fund social programs, they are similarly resentful. When *I* appeal to morality, I'm just speaking up for doing the right and decent thing, but when *someone else* tells me to do something I don't want to do in the name of "morality," he's at it again, trying to ram *his* morality down *my* throat.

Moral realists can maintain a distinction between morality and moralizing. Being moral is getting morality correct, while moralizing is getting morality wrong, either by upholding false moral demands or, what amounts to the same thing, overextending the principles of true morality. Without moral realism, however, there is no objective difference between morality and moralism, just competing values. Naturally, some people's values place greater demands on others, and one might hope to separate the moral from the moralistic based on the extent to which different people's values demand interference with other people's lives and decisions. But by that measure, anarchists are the least moralistic of all.

The solution to the tragedy of common sense morality lies in our recognizing that moralism and morality are the same thing seen from two different vantage points. We must extend some of our distrust of moralism to morality, while at the same time extending some of our natural affinity for morality to moralism. In other words, we must recognize that our own moral intuitions are merely reflections of our own values and be willing to compromise them when necessary, and at the same time we must not dismiss other people's values simply because they are at odds with our own. We must be a bit more detached

and cynical when it comes to our own values, and a bit more open-minded and welcoming when it comes to the values of others. Calls for tolerance, of course, are nothing new, but the revisionist proposal is more than just a *call* for tolerance. It's a program for sowing the seeds of tolerance by understanding and counteracting the psychological mechanisms and philosophical presuppositions that make tolerance difficult.

So that's it. Problem understood, problem solved. We'll adopt revisionism and enjoy the social benefits of thought and speech that reflect the nature of morality. Were it only that easy! Getting a handful of philosophers to appreciate the tragedy of common sense morality and phase out their realist tendencies is one thing. Getting the general public to appreciate the problem and commit to its solution is another. In fact, the tragedy of common sense morality is really *two* tragedies. Before we can solve the main problem of getting people to curb their realism in order to avoid its costs, we have to solve *the problem of getting people to see that there's a problem*, the problem of getting people to see that moral realism *has* costs. Getting people to see, for example, that the herdspeople are raising too many cows on the pasture or that the river is dangerously polluted is *relatively easy*. In such cases, many will readily acknowledge the problem; the hard part is working out a solution to which everyone can agree. In the case of moral realism, however, the connection between the tempting behavior and the harm it causes is far less clear. People know that the world is full of suffering and conflict, but they have no idea that common sense morality has anything to do with it. On the contrary, most people think that in treating the world's social

ills we need *more* good-old-fashioned, common sense morality, not *less*. That common sense morality is a good thing is *common sense*—meta-common sense, if you will—and fighting against common sense, meta or otherwise, is notoriously difficult, all the more so in the moral domain. What to do?

Getting large numbers of people to put aside common sense isn't impossible. As you skid down an icy road, your untutored physical intuitions will tell you to turn *out* of the skid, but millions of informed drivers know that the best thing to do, contrary to common sense, is to turn *into* the skid. Why was this bit of anti-common sense able to gain such widespread acceptance? To put the question in Dawkinsian terms: What accounts for the success of the "turn into the skid" meme?[47] It's likely that its success has something to do with the fact that this meme directly benefits its "host." People are willing to pay attention to automotive safety tips because they stand to benefit greatly and directly from doing so. Government officials are willing to force people to pay attention to automotive safety tips because the benefits are relatively easy to demonstrate empirically. Unfortunately, the revisionism meme and the "turn into the skid" meme are dissimilar in these regards. The revisionist meme doesn't benefit its host directly, and may not benefit its host at all, may even *harm* its host, if there

---

[47] A meme is, roughly, an idea, viewed as analogous to a gene. A meme can be anything from a catchy tune or a joke to a design for an atomic bomb. Memes spread by jumping from host to host, i.e. from brain to brain, and they replicate or fail to replicate in virtue of their functional properties. See Dawkins (1976) for the seminal discussion. See Blackmore (1999) for a recent popular account.

aren't enough copies of the revisionist meme in circulation, which there are not at this time. Another important difference between these two memes is that the revisionist meme is relatively expensive to acquire. Getting the revisionist meme would appear to involve something like reading this essay, finding it convincing, and forming an intention to further the revisionist cause. But that's just the home stretch. Having the motivation, patience, and ability to read and appreciate long philosophical essays tends to require a lot of education, a lot more than most people get, and may require a relatively rare sort of curiosity. Moreover, when it comes to philosophical reading, people are drawn to material with which they expect to agree, material that is likely to enrich one's current worldviews rather than change it,[48] especially when it comes to moral values. Thus, spreading the revisionist meme appears to require getting people, many of whom lack the requisite education, to devote their time and energy to appreciating complicated

---

[48] I have a distinct memory of my first encounter with John Mackie's Book *Ethics*: *Inventing Right and Wrong*. I was a sophomore philosophy major, interested in ethics, but thoroughly realist in my thinking. I was making one of my regular visits to the small shelf devoted to used philosophy books in the basement of the Harvard Book Store. The main title "Ethics" caught my eye. I picked it up, read the subheading, "Inventing Right and Wrong," and promptly returned it to the shelf. I didn't read Mackie's book until some years later when I began to have doubts about moral realism and found the standard non-cognitivist alternatives unsatisfying.

philosophical/scientific ideas that are not going to benefit them and that fly in the face of common sense values. This does not sound promising.

But all hope is not lost. There are, I think, two approaches to spreading revisionism, the "hard sell" and the "soft sell." What I described above is essentially the hard sell, saying to people, "Here! Read this! It explains how there's no fact of the matter about what's right or wrong and how our brains trick us into thinking that there is and how we can make the world better by not using words like 'right' and 'wrong.' Of course, reading this won't make you happier or richer or better looking, but you really should read it anyway because your reading it will make it easier for people to get along with you!" Needless to say, this strategy is not likely to meet with much success. People everywhere like the values they've got just fine—it's *other people's* values that are the problem—and they're not going to do a lick of extra work to revise their moral outlooks.[49] As noted above (Section 5.1.4), people don't seek out factual knowledge in hopes that that knowledge will change their values.[50] Rather, they seek out factual knowledge for other reasons, because it's useful or inherently interesting, and then find, sometimes, that their new knowledge has somehow brought about a change in evaluative perspective.

---

[49] What about philosophers? It seems to me that moral philosophers are not so much interested in changing their moral intuitions as they are in organizing them and justifying them, but perhaps this is not always the case.

[50] By "values" here I mean things people care about for their own sakes and not for instrumental reasons.

Consider someone who has the common sense intuition that the incestuous relationship between Mark and Julie (See Section 3.2) is absolutely, positively *wrong*. Suppose you say to her, "Here! Read this book on evolutionary psychology. It offers biologically based explanations of many behavioral phenomena, including the incest taboo. When you're done reading it you may not be so sure that all incest is wrong." This is a good way to convince someone *not* to read your book. But suppose instead you say, "Here. You might find this book on evolutionary psychology interesting. It offers biologically based explanations of many important aspects of the human experience, from love to violence to morality." Now you've got a chance. The topics are inherently interesting, and the arguments are compelling once you get into them. It's not hard to see how a common sense moralist with a bit of curiosity about human nature would pick up a book on evolutionary psychology, and it's not hard to see how someone who did so could emerge with doubts about her old values, some of them at least.   This is the soft sell. Pushing revisionist ideas directly is unlikely to meet with much success.

In light of this, I've come to the following conclusion: The most important thing we can do to further the revisionist cause is to further the advancement of the science of human nature. And this science is booming.[51] Darwin's theory of natural selection (1964) provides the over-arching framework. The fields of cognitive neuroscience and behavioral genetics are uncovering the biological mechanisms that make us who we are. Cognitive psychology continues to

---

[51] Summary follows Pinker (2002).

further our understanding of the human mind as an information processor.

Economics, anthropology, and sociology offer a broader perspective. These disciplines are becoming increasingly interdependent, forming a unified science of human nature that is as powerful as it is profound (Wilson, 1998). Emerging from this global enterprise are countless hypotheses concerning every aspect of human experience, from the depths of disease to the heights of religious experience, but, more than that, this enterprise has spawned a *worldview*, the so-called "scientific worldview." It is no doubt misleading to speak of *the* scientific worldview, as there are many, but they are, I think, more similar than dissimilar. The scientific worldview is not new, but with each passing decade it is broadened and refined. Much silliness and nastiness has been spoken and justified in the name of science and, by extension, its worldview, but that is to be expected. The better an idea is, the more tempting its abuse. In large part due to its historical abuse, the scientific worldview has received much resistance from the left, and speaking on its behalf smacks of naïveté in many intellectual circles. But the left is slowly coming around. These days, the strongest opposition is from the right, particularly the religious right. Their opposition remains fierce, and for good reason. Religion has a variety of commitments that continue to be undermined by science, both directly in the case of factual commitments and indirectly in the case of evaluative ones. The public at large remains relatively neutral between these clashing worldviews, favoring one over the other depending on the context, although at this time, at least in the United States, religion appears to hold sway.

People who reject the scientific worldview are highly unlikely to embrace revisionism. In contrast, people who embrace the scientific worldview and, more specifically, people for whom evolution provides the framework for their understanding of human nature,[52] are revisionists waiting to happen. Thinkers for whom evolution looms large have, I think, been a bit uneasy when it comes to normative ethics. Their understanding of the bio-cultural bases for moral thought and behavior makes them partial moral skeptics, i.e. *moral$_1$* skeptics. Moral intuitions that seem like perceptions of an external moral reality can no longer be understood as such. But at the same time, evolutionary thinkers are still humans with human values. They still care about the well-being of others and take our efforts to make the world a better place seriously. Hence, they are not complete moral skeptics because they are not *moral$_2$* skeptics. The meta-ethical unease experienced by evolutionary thinkers is, in my opinion, the result of an inadequacy in our ordinary moral concepts, which is reflected in ordinary moral language. The evolutionarily minded don't know what to think about "morality" because they've yet to distinguish explicitly between the scientifically implausible *morality$_1$* and the ever-so-important *morality$_2$*. As soon as one realizes that morality$_1$ and morality$_2$ are separable and that the latter does not depend on the former, revisionism is a natural position to adopt.

---

[52] In my experience, most contemporary ethicists working in the Anglo-American philosophical tradition accept the theory of evolution, but evolution is not central to their thinking. There are, however, some exceptions, and their ranks are growing. See, for example, Singer (2000) and Joyce (2001).

The evolutionary perspective makes one curious about and receptive to information concerning moral psychology. The projectivist account of realist psychology and phenomenology helps clear away lingering realist doubts, and the social intuitionist account of moral development explains why moral disagreement is so frustrating and recalcitrant. Further reflection on the nature of moral psychology and its likely effects in our present context suggest that $moral_1$ language and thought are not only unnecessary but counterproductive, and thus getting rid of them starts to seem like an appealing idea. At the same time, an understanding of moral intuition and the rejection of moral realism make deontological line-drawing seem artificial and make consequentialism more appealing by default. The suggestion that utilitarian concerns are recognized as legitimate *across moral divides* and that therefore utilitarianism is a shareable moral framework furthers its appeal.

All told, revisionism is a natural fit for those who have appreciated what science has to say about the nature of morality, and as a result one need not fight too hard to promote revisionism directly. The key to moral progress most likely lies in the advancement of the science of human nature.

5.5 Saving the World

What are the biggest problems we humans face? The list of overlapping topics looks something like this: the proliferation of weapons of mass destruction, the

destruction of the environment, ethnopolitical conflict, terrorism, overpopulation, infectious disease, hunger, and malnutrition. What these problems have in common is that they are preventable by, if not directly caused by, human choice. Solving these problems requires compromise, but these compromises have, for the most part, yet to be made. Why? Is it just selfishness? Do people simply not care? That may be the case for some, but I'm convinced that most people and most leaders care about the future of the world and believe they are doing what is right. The problem, of course, is that everyone has a different conception of what is right, of what is required of themselves and of others. When different people's respective visions for the future fail to dovetail, well-meaning people throw up their hands and say, "What can I do? They leave me no choice."

The looming problems named above are all variations on Hardin's tragedy of the commons, and the problem that is the subject of this essay, the problem of transcending common sense morality, is the mother of all commons problems. Once again, commons problems have two kinds of solutions: political and moral. Since World War I, globally minded people have dreamed of a political solution to the world's most pressing problems in the form of a world government. Perhaps a single coalition will take over the world by force and from its position of unprecedented power impose upon humanity some much-needed discipline. This is unlikely to happen, and were the world to be united by force the cure would probably be as bad as the disease. More likely, the nations of the world will have to choose to surrender their powers voluntarily. And that means that the favored political solution to our global woes is ultimately a moral solution, one

that must be enacted voluntarily. Could that happen? Not as long as the people

of the world lack a common framework for resolving differences across

ideological divides. High minded words such as "rights," "freedom," and "justice"

are not enough because these words mean different things to different people.

We need a common currency for moral exchange, a mutually agreeable system

that can resolve moral questions rather than begging them. But we will never

have such a system so long as people hold fast to their moral intuitions, taking

them to be messengers of moral truth. We must understand where our intuitions

come from and why they vary so widely.

Once again, we are saddled with a Stone Age moral psychology that is

appropriate to life in small, homogeneous communities in which all members

share roughly the same moral outlook. Our minds trick us into thinking that *we*

are *absolutely right* and that *they* are *absolutely wrong* because, once upon a

time, this was a useful way to think. It is no more, though it remains natural as

ever. We love our respective moral senses. They are as much a part of us as

anything. But if we are to live together in the world we have created for

ourselves, so unlike the one in which our ancestors evolved, we must know when

to trust our moral senses and when to ignore them.

I conclude with the words of two leaders, one wise, one otherwise.

"States like these, and their terrorist allies, constitute an axis of evil,"[53]

—George W. Bush

---

[53] From the State of the Union Address, 2002.

"If only it were so simple! If only there were evil people somewhere insidiously committing evil deeds, and it were necessary only to separate them from the rest of us and destroy them. But the line dividing good and evil cuts through the heart of every human being. And who is willing to destroy a piece of his own heart?"[54]

—Aleksandr Solzhenitsyn

---

[54] From *The Gulag Archipelago* (1974).

Bibliography

Anderson, S. W., A. Bechara, H. Damasio, D. Tranel, and A. R. Damasio. (1999). "Impairment of social and moral behavior related to early damage in human prefrontal cortex," *Nature Neuroscience*, 2 (11), 1032-7.

Anscombe, G. E. M. (1958). "Modern Moral Philosophy," *Journal of the Royal Institute of Philosophy*, 33 (124), 1-19.

Aristotle (1985). *Nicomachean Ethics*, Irwin, T. trans. Indianapolis: Hackett Publishing Co.

Armstrong, D. M. (1998). "Truths and Truthmakers." unpublished manuscript, Dept. of Philosophy, Sydney University.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Ayer, A. J. (1952). *Language, Truth, and Logic*. New York: Dover.

Baron, J. (1998). *Judgment Misguided*. New York: Oxford University Press.

Bentham, Jeremy (1982). *An Introduction to the Principles of Morals and Legislation*, J. H. Burns and H. L. A. Hart, eds.. London, New York: Methuen. (Originally published, 1789.)

Berry, D. S. and K. Springer. (1993). "Structure, Motion, and Preschoolers' Perception of Social Causality," *Ecological Psychology*, 5, 273-283.

Blackburn, S. (1984). *Spreading the Word*. New York: Oxford University Press.

Blackburn, S. (1993a). *Essays in Quasi-Realism*. New York: Oxford University Press.

Blackburn, S. (1993b). "Realism, Quasi or Queasy?" in J. Haldane. and C. Wright, eds., *Reality, Representation, and Projection*. New York: Oxford University Press.

Blackmore, S. (1999). *The Meme Machine*. New York: Oxford University Press.

Blair, R. J. R. (1995), "A Cognitive Developmental Approach to Morality: Investigating the Psychopath," *Cognition*, 57, 1-29.

Blair, R. J. R., L. Jones, F. Clark, and M. Smith (1997). "The Psychopathic Individual: A Lack of Responsiveness to Distress Cues?" *Psychophysiology*, 34, 192-198.

Boyd, R. (1988). "How to Be a Moral Realist," in G. Sayre-McCord, ed. *Essays on Moral Realism*. Ithaca: Cornell University Press.

Brandt, R. B. (1979). *A Theory of the Good and the Right*. Oxford: Clarendon Press.

Burgess, J. A. (1998). "Error Theories and Values," *Australasian Journal of Philosophy*, 76 (4), 534-552.

Burgess, J. P. (1979). "Against Ethics," unpublished manuscript, Dept. of Philosophy, Princeton University.

Calaprice, A., ed. (2000). *The Expanded Quotable Einstein*. Princeton: Princeton University Press.

Chalmers, D. (1995). "Facing Up to the Problem of Consciousness," *Journal of Consciousness Studies*, 2 (3), 1995, pp. 200-219

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Coen, E. and J. Coen, (1998). *The Big Lebowski*. London: Faber and Faber.

Cohen, R. (1991). *Negotiating Across Cultures: Communication Obstacles in International Diplomacy*. Washington, D.C.: United States Institute of Peace.

Csibra, G., G. Gergely, S. Biro, O. Koos. and M. Brockbank, (1999). "Goal Attribution Without Agency Cues: The Perception of 'Pure Reason' in Infancy," *Cognition*, 72, 237-267.

Dehaene, S. (1997). *Number Sense: How the Brain Creates Mathematics*. New York: Oxford University Press.

Damasio, A. R., D. Tranel, and H. Damasio, (1990). "Individuals with Sociopathic Behavior Caused by Frontal Damage Fail to Respond Autonomically to Social Stimuli," *Behavioral Brain Research*, 41, 83-94.

Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Grosset/Putnam.

Damasio, A. (1999). *The Feeling of What Happens*. New York: Harcourt Brace.

Darwall, S., A. Gibbard, and P. Railton (1992). "Toward a Fin de Siecle Ethics: Some Trends," *Philosophical Review*, 101, 115-89.

Darwin, C. (1964). *On the Origin of Species*. Cambridge, MA: Harvard University Press. (Originally published, 1859.)

Davidson R. J., and W. Irwin (1999). "The Functional Neuroanatomy of Emotion and Affective Style." *Trends In Cognitive Sciences*, 3 (1), 11-21.

Davies, M. and L. Humberstone, (1980). "Two Notions of Necessity," *Philosophical Studies*, 38, 1-30.

Dawkins, R. (1976). *The Selfish Gene*. New York: Oxford University Press.

Dennett, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Schuster.

de Waal, F. (1982). *Chimpanzee Politics: Power and Sex Among the Apes*. New York: Harper & Row.

de Waal, F. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

Fehr, E. and S. Gachter (2002). "Altruistic Punishment in Humans," *Nature* 415, 137-40.

Fisher, R. (1971). *Basic Negotiating Strategy: International Conflict for Beginners*. London: Allen Lane.

Fisher, R. (1994). *Beyond Machiavelli: Tools for Coping with Conflict*. Cambridge, MA: Harvard University Press.

Fishkin, J. S. (1984). *Beyond Subjective Morality: Ethical Reasoning and Political Philosophy*. New Haven, Yale University Press.

Fiske, A. P. (1991). *Structures of Social Life*. New York: Free Press.

Fiske, A. (1999). "Learning Culture the Way Informants Do: Observing, Imitating, and Participating." University of California, Los Angeles.

Foot, P. (1972), "Morality as a System of Hypothetical Imperatives," *Philosophical Review*, 71, 305-316.

Foot, P. (1994). "Recantation 1994," in Michael Smith, ed. *Meta-Ethics*. Aldershot: Dartmouth.

Frankfurt, H. G. (1971). "Freedom of the Will and the Concept of a Person," *Journal of Philosophy*, 68, 5-20.

Garner, R. (1994). *Beyond Morality*. Philadelphia: Temple University Press.

Gast L., and T. Hackford, (1997). *When We Were Kings*. New York: Polygram.

Gazzaniga, M. S. (1985). *The Social Brain*. New York: Basic Books.

Geach, P. T. (1965). "Assertion," *Philosophical Review*, 74, 449-465.

Gergely, G., Z. Nadasdy, G. Csibra, and S. Biro (1995). "Taking the Intentional Stance at Twelve Months of Age," *Cognition*, 56, 165-193.

Gettier, E. L. (1963). "Is Justified True Belief Knowledge?" *Analysis*, 23, 121-3.

Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.

Goodin, R. E. (1995). *Utilitarianism as a Public Philosophy*. New York: Cambridge University Press.

Greene, J. D. (2001). "A Psychological Perspective on Nozick's Experience Machine and Parfit's Repugnant Conclusion." Presentation at the Society for Philosophy and Psychology Annual Meeting, Cincinnati, OH.

Greene, J. and J. Baron (2001). "Intuitions About Declining Marginal Utility," *Journal of Behavioral Decision Making*, 14 (3), 243-55.

Greene J. D., R. B. Sommerville, L. E. Nystrom, J. M. Darley, and J. D. Cohen (2001). "An fMRI Investigation of Emotional Engagement in Moral Judgment" *Science*, 293 (5537), 2105-8.

Haidt, J. (2000). unpublished manuscript, Dept. of Psychology, University of Virginia. Published in shortened form as Haidt (2001).

Haidt, J. (2001). "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review*, 108 (4), 814-834.

Haidt, J., F. Bjorklund, and S. Murphy (2000). "Moral Dumbfounding: When Intuition Finds no Reason," unpublished manuscript, Dept. of Psychology, University of Virginia.

Haidt, J. and M. Hersh, (2001). "Sexual Morality: The Cultures and Reasons of Liberals and Conservatives," *Journal of Applied Social Psychology*, 31, 191-221.

Hardin, C. L. (1988). *Color for Philosophers: Unweaving the Rainbow*. Indianapolis: Hackett Pub. Co.

Hardin, G. (1968). "The Tragedy of the Commons," *Science*, 162, 1243-1248.

Hare, R. D. and M. J. Quinn (1971). "Psychopathy and Autonomic Conditioning," Journal of Abnormal Psychology, 77 (3), 223-35.

Hare, R. M. (1952). *The Language of Morals*. New York: Oxford University Press.

Hare, R. M. (1981). *Moral Thinking: Its levels, method, and point*. New York: Oxford University Press, 1981.

Harman, G. (1977). *The Nature of Morality*. New York: Oxford University Press.

Harman, G. and J. J. Thomson (1996). *Moral Relativism and Moral Objectivity*. Cambridge: Blackwell.

Harris, P. L. (1989). *Children and Emotion*. Oxford: Basil Blackwell.

Harsanyi, J. C. (1982), "Morality and the Theory of Rational Behaviour," in A. Sen. And B. Williams, eds., *Utilitarianism and Beyond*. New York: Cambridge University Press.

Hashimoto, H. (1966). "A Phenomenal Analysis of Social Perception," *Journal of Child Development*, 2, 3-26.

Hausfater, G. and S. B. Hrdy, eds., (1984). *Infanticide: Comparative and Evolutionary Perspectives*. New York: Aldine..

Heider, F. and M. Simmel (1944). "An Experimental Study of Apparent Behavior," *American Journal of Psychology*, 57, 243-249.

Hinckfuss, I. (1987), "The Moral Society: Its Structure and Effects," Preprint Series in Environmental Philosophy, Department of Philosophy, 16, Australian National University.

Horgan, T. and M. Timmons (1992). "Trouble on Moral Twin Earth: Moral Queerness Revived," *Synthese*, 92 (2), 221-260.

Hume, D. (1978). *A Treatise of Human Nature*, L. A. Selby-Bigge and P. H. Nidditch, eds. Oxford: Clarendon. (Originally published, 1739)

Hume, D. (1960). *An Enquiry Concerning the Principles of Morals*. La Salle, IL: Open Court. (Originally published, 1777)

Huttenlocher, P. R. (1994). "Synaptogenesis, Synapse Elimination, and Neural Plasticity in Human Cerebral Cortex," in C. A. Nelson, ed., *Threats to Optimal Development: Integrating Biological, Psychological, and Social Risk Factors*. The Minnesota Symposia on Child Development, 27., 35-54. Hillsdale, NJ: Lawrence Erlbaum.

Jackson, F. (1974). "Defining the Autonomy of Ethics," *Philosophical Review*, 83, 88-96.

Jackson, F. (1998). *From Metaphysics to Ethics*. Oxford: Clarendon Press.

Johnston, M. (1989), "Dispositional Theories of Value," *Proceeding of the Aristotelian Society*, Supp. vol. 63, 113-37.

Joyce, R. (2001). *The Myth of Morality*. Cambridge: Cambridge University Press.

Kant, I (1993). *Grounding for the Metaphysics of Morals; with "On a Supposed Right to Live Because of Philanthropic Concerns,"* J. W. Ellington, trans., 3rd ed.. Indianapolis: Hackett Publishing Co..

Kanger, S. and H. Kanger (1966). "Rights and Parliamentarianism," *Theoria*, 32, 85-115.

Kennan, G. F. (1985), "Morality and Foreign Policy," *Foreign Affairs*, 64 (2), 205-218.

Kohlberg, L. (1969). "Stage and Sequence: The Cognitive-Developmental Approach to Socialization," in Goslin, D. A., ed. *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.

Kohlberg, L. (1984). *The Psychology of Moral Development : the Nature and Validity of Moral Stages*. San Francisco: Harper & Row.

Korsgaard, C. (1996). *The Sources of Normativity*. New York: Cambridge University Press.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Cambridge, MA: Harvard University Press.

Kuhn, Deanna (1991). *The Skills of Argument*. New York: Cambridge University Press.

Kunst-Wilson, W. R. and R. B. Zajonc (1980). "Affective Discrimination of Stimuli that Cannot Be Recognized," *Science*, 207, 557-558.

Kurtzman, D. R. (1970). "'Is,' 'Ought,' and the Autonomy of Ethics," *Philosophical Review*, 79, 493-509.

Lakoff, G. (1996). *Moral Politics: What Conservatives Know that Liberals Don't*. Chicago: University of Chicago Press.

Lakoff, G. and M. Johnson (1999). *Philosophy in the Flesh*. New York: Basic Books.

Lewis, D. K. (1970). "How to Define Theoretical Terms," *Journal of Philosophy*, 67, 427-46.

Lewis, D. K. (1989). "Dispositional Theories of Value," *Proceeding of the Aristotelian Society*, Supp. vol. 63, 113-37.

Lieberman, M. D. (2000). "Intuition: A Social Cognitive Neuroscience Approach," *Psychological Bulletin*, 126, 109-137.

Llinas, R. (2001). *I of the Vortex*. Cambridge, MA: MIT Press.

MacIntyre, A. (1981). *After Virtue*. Notre Dame: Notre Dame University Press.

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. New York: Penguin.

Maddock, R. J. (1999). "The Retrosplenial Cortex and Emotion: New Insights from Functional Neuroimaging of the Human Brain," *Trends in Neuroscience*, 22 (7), 310-6.

Mahapatra, M. (1981). *Traditional Structure and Change in an Orissan Temple*. Calcutta: Punthi Pustak.

Maier, N. R. F. (1931), "Reasoning in Humans: II. The Solution of a Problem and Its Appearance in Consciousness," *Journal of Comparative Psychology*, 12, 181-194.

McCloskey, M. (1983). "Intuitive Physics," Scientific American 248, 122-130.

McDowell, J. (1985). "Values and Secondary Properties," in T. Honderich, ed. *Morality and Objectivity*. London: Routledge and Kegan Paul.

McNaughton, D. (1988). *Moral Vision: An Introduction to Ethics*. New York: Blackwell.

Mill, J. S. (1998). *Utilitarianism*, Crisp. R., ed.. New York: Oxford University Press. (Originally published, 1861.)

Minsky, M. (1985). *The Society of Mind*. New York: Simon and Schuster.

Minoura, Y. (1992). "A Sensitive Period for the Incorporation of Cultural Meaning System: A Study of Japanese Children Growing Up in the United Sates," *Ethos*, 20, 304-339.

Montville, J. (1993). "The Healing Function in Political Conflict," in D. J. D. Sandole and H. van der Merwe, eds., *Conflict Resolution Theory and Practice: Integration and Application*. Manchester: Kelman Publication.

Moore, G. E.. (1959). *Principia Ethica*. Cambridge: Cambridge University Press. (Originally published, 1903.)

Morris, M. W. and K. Peng (1994). "Culture and Cause: American and Chinese Attributions for Social and Physical Events," *Journal of Personality and Social Psychology*, 67, 949-971.

Neurath, O. (1959). "Protocol Sentences," in A. J. Ayer, ed. *Logical Positivism*. New York: Free Press.

Nisbett, R. E. and T. D. Wilson (1977), "Telling More than We Know: Verbal Reports on Mental Processes," *Psychological Review*, 84, 231-259.

Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.

Nozick, R. (1993). *The Nature of Rationality*. Princeton, NJ: Princeton University Press.

Nussbaum, M. C. (1995). *Poetic Justice: The Literary Imagination and Public Life*. Boston, MA: Beacon Press.

Parfit, D. (1984). *Reasons and Persons*. Oxford: Clarendon Press.

Pinker, S. (1994). *The Language Instinct*. New York: W. Morrow and Co.

Pinker, S. (2002). *The Blank Slate*. Unpublished manuscript, Dept. of Brain and Cognitive Sciences, M.I.T..

Prior, A. N. (1960). "The Autonomy of Ethics," *Australasian Journal of Philosophy*, 38, 199-206.

Raine, R., T. Lencz, B. Bihrle, L. LaCasse, and P. Colletti (2000). "Reduced Prefrontal Gray Matter Volume and Reduced Autonomic Activity in Antisocial Personality Disorder," *Archives of General Psychiatry*, 57, 119-27.

Raine, A., M. S. Buchsbaum, J. Stanley, S. Lottenberg, L. Abel, and J. Stoddard (1994). "Selective Reduction in Prefrontal Glucose Metabolism in Murderers," *Biological Psychiatry*, 36, 365-373.

Ramachandran, V. S. and E. M. Hubbard (2001). "Synaesthesia — A Window Into Perception, Thought and Language," *Journal of Consciousness Studies*, 8 (12).

Rao, R. V. R. C. (1976). "Searching for a Mature Relationship," *Round Table*, 263, 249-60.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rawls, J. (1980). "Construction and Objectivity," *Journal of Philosophy*, 77, 554-72.

Reiman, E. M. (1997). "The Application of Positron Emission Tomography to the Study of Normal and Pathological Emotions," *Journal of Clinical Psychiatry*, 58 (supp. 16), 4-12.

Rime, B. *et al*. (1985). "The Perception of Interpersonal Emotions Originated by Patterns of Movement," *Motivation and Emotion*, 9, 241-260.

Robinson, R. J., D. Keltner, A. Ward, L. Ross (1995). "Actual Versus Assumed Differences in Construal: "Naïve Realism" in Intergroup Perception and Conflict," *Journal of Personality and Social Psychology,* 68, 404-417.

Rozin, P., A. E. Fallon, and M. Augustoni-Ziskind (1985). "The Child's Conception of Food: The Development of Contamination Sensitivity to "Disgusting" Substances," *Developmental Psychology*, 21, 1075-1079.

Rozin, P., L. Lowery, J. Haidt, and S. Imada (1999). "The CAD Triad Hypothesis: A Mapping Between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, and Divinity)," *Journal of Personality and Social Psychology*, 76 (4), 574-586.

Russell, B. (1946). "Reply to Criticisms," in P. A. Schilpp, ed., *The Philosophy of Bertrand Russell*. Evanston: Library of Living Philosophers, Inc.

Saver, J. L. and A. R. Damasio, (1991). "Preserved Access and Processing of Social Knowledge in a Patient with Acquired Sociopathy Due to Ventromedial Frontal Damage," *Neuropsychologia*, 29 (12), 1241-1249.

Sayre-McCord, G. (1988). "The Many Moral Realisms" in G. Sayre-McCord, ed. *Essays on Moral Realism*. Ithaca: Cornell University Press.

Scanlon, T. M. (1982). "Contractualism and Utilitarianism," in A. Sen and B. Williams, eds., *Utilitarianism and Beyond*. New York: Cambridge University Press.

Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Schelling, T. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Schlesinger, A. (1971). "The Necessity of Amorality in Foreign Affairs," *Harper's Magazine*, August, 1971.

Scholl, B. J. and P. D. Tremoulet, (2000). "Perceptual Causality and Animacy," *Trends in Cognitive Sciences*, 4 (8), 299-309.

Shore, B. (1996). *Culture in Mind: Cognition, Culture, and the Problem of Meaning*. New York: Oxford University Press.

Shweder, R. A., M. Mahapatra and J. Miller (1987). "Culture and Moral Development," in J. Kagan and S. Lamb, eds., *The Emergence of Morality in Young Children*. Chicago: University of Chicago Press.

Shweder, R. A., N. C. Much, M. Mahapatra, and L. Park (1997). "The 'Big Three' of Morality (Autonomy, Community, and Divinity) and the 'Big Three' Explanations of Human Suffering," in A. Brandt and P. Rozin, eds., *Morality and Health*. New York: Routledge.

Shweder, R. A., J. Goodnow, G. Hatano, R. A. LeVine, H. Markus, and P. Miller (1998). "The Cultural Psychology of Development: One mind, Many Mentalities," in W. Damon, ed. *Handbook of Child Psychology*, 5th Ed., Vol. 1, R. M. Lerner, ed., *Theoretical Models of Human Development*. New York: Wiley.

Sidgwick, H. (1890). *The Methods of Ethics*, 4th ed.. London: Macmillan.

Singer, P. (1972). "Famine, Affluence, and Morality," *Philosophy and Public Affairs*, 1, 229-243.

Singer, P. (1975). *Animal Liberation : A New Ethics For Our Treatment of Animals*. New York: New York Review (Distributed by Random House).

Singer, P. (1979). *Practical Ethics*. New York: Cambridge University Press.

Singer, P. (1995). *Rethinking Life and Death: The Collapse of Traditional Ethics*. New York: St. Martin's Press.

Singer, P. (2000). *A Darwinian Left: Politics, Evolution, and Cooperation*. New Haven: Yale University Press.

Smart, J. J. C. and B. Williams (1973). *Utilitarianism For and Against*. Cambridge: Cambridge University Press.

Smart, J. J. C. (1961). *An Outline of a System of Utilitarian Ethics*. Melbourne: Melbourne University Press.

Sober, E. and D. S. Wilson (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*.

Solzhenitsyn, A. I. (1974). *The Gulag Archipelago, 1918-1956; an Experiment in Literary Investigation*, Whitney, T. P., trans.. New York: Harper and Row.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Stevenson, C. L. (1944). *Ethics and Language*. New Haven: Yale University Press.

Stirner, M. (1973). *The Ego and His Own*. New York: Dover.

Sturgeon, N. (1988). "Moral Explanations" in G. Sayre-McCord, ed. *Essays on Moral Realism*. Ithaca: Cornell University Press.

Thompson, K. (1985). *Moralism and Morality in Politics and Diplomacy*. Lanham: University Press of America.

Thomson, J. J. (1986). *Rights, Restitution, and Risk*. Cambridge, MA: Harvard University Press.

Thomson, J. J. (1990). *The Realm of Rights*. Cambridge, MA: Harvard University Press.

Timmons, M. (1999). *Morality Without Foundations: A Defense of Ethical Contextualism*. New York: Oxford University Press.

Trainer, F. E. (1982). *Dimensions of Moral Thought*. Sydney: New South Wales University Press.

Unger, P. (1996). *Living High and Letting Die: Our Illusion of Innocence*. New York: Oxford University Press.

Viorst, J. (1987). *Alexander and the Terrible, Horrible, No Good, Very Bad Day*, reissue edition. New York: Simon and Schuster.

Warnock, G. J. (1971). *The Object of Morality*. London: Methuen.

Werker, J. and R. Tees (1984). "Cross-language Speech Perception: Evidence for Perceptual Reorganization During the First Year of Life," *Infant Behavior and Development*, 7, 49-63.

Williams, B. (1985a). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Williams, B. (1985b). "Ethics and the Fabric of the World," in T. Honderich, ed. *Morality and Objectivity*. London: Routledge and Kegan Paul.

Wilson, E. O. (1998). *Consilience*. New York: Knopf.

Wiskott, L., J.-M. Fellous, N. Kruger, and C. von der Marlsburg (1995). "Face Recognition and Gender Determination," *Proceedings of the International Workshop on Automatic Face- and Gesture recognition*. Zurich.

Wright, R. (1994). *The Moral Animal: Evolutionary Psychology and Everyday Life*. New York: Pantheon.

Wright, R. (2000). *Nonzero: The Logic of Human Destiny*. New York: Pantheon.

Zajonc, R. B. (1984). "On the Primacy of Affect," *American Psychologist*, 39 (2), 117-123.

Zimbardo, P. G., S. LaBerge, and L. D. Butler (1993). "Psychological Consequences of Unexplained Arousal: A Posthypnotic Suggestion Paradigm," *Journal of Abnormal Psychology*, 102, 466-473.