INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

THE UNIVERSITY OF CHICAGO

ADVICE, LIFE-EXPERIENCE, AND MORAL OBJECTIVITY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE HUMANITIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHILOSOPHY

BY

ERIC WILAND

CHICAGO, ILLINOIS

AUGUST 1997

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 9800665

UMI Microform 9800665
Copyright 1997, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized
copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table of Contents

ACKNOWLEDGMENTS v

ABSTRACT vii

CHAPTER ONE. ADVICE AND ACTION 1

1.1 REASONS FOR ACTION AND REASONS FOR BELIEF 1
1.2 TESTIMONY 3
1.3 THE POSSIBILITY OF ADVICE 7
1.4 ADVICE AND KNOWING WHAT TO DO 15
1.5 SOME DEAD ENDS 23
1.6 LIFE-EXPERIENCE: DOING WELL 26
1.7 LIFE-EXPERIENCE: NOT DOING SO WELL 31
1.8 CONCLUSION 40

CHAPTER TWO. UNDERSTANDING MORAL OBJECTIVITY 42

2.1 METAETHICS AND PHILOSOPHICAL PSYCHOLOGY 42
2.2 TWO THOUGHTS ABOUT OBJECTIVITY 43
2.3 EXPLAINING MORAL VIEWS 50
2.4 OBJECTIVITY, CAUSATION AND COUNTERFACTUALS 56
2.5 OBJECTIVITY AND JUSTIFICATION 72
2.6 THE OBJECTIVITY OF ARTIFACTS 74
2.7 THE ABSOLUTE CONCEPTION 77
2.8 THE NEED FOR A THEORY OF ERROR 90
2.9 CONCLUSION 100

CHAPTER THREE. THE INS AND OUTS OF REASONS 102

3.1 INTRODUCTION 102
3.2 EXPLANATION AND JUSTIFICATION 104
3.3 WILLIAMS' CORRECTIONS TO THE SUB-HUMEAN MODEL 108
3.4 WILLIAMS' ARGUMENT AGAINST EXTERNAL REASONS THEORY 112
3.5 THE VACUITY OF WILLIAMS' ARGUMENT 118
3.6 SEEKING NORMATIVITY 122
3.7 MIXING MINDS AND THE LIMITS OF PHILOSOPHICAL SURGERY 131
3.8 IMPLICATIONS 136

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER FOUR. ARE THE IMMORAL IGNORANT? 138

4.1 Lacking information 138
4.2 Bad philosophy 142
4.3 Looking back at advice 150
4.4 Experience and literature 158
4.5 Autobiography as advice 163
4.6 Reacting to our inexperience: some objections 169
4.7 Conclusion 178

CHAPTER FIVE. DO THE IMMORAL LACK INTEGRITY? 180

5.1 Not knowing, not caring 180
5.2 Collecting your thoughts 182
5.3 Looking for the lukewarm 184
5.4 Fragmented Nazis 189
5.5 Repressing humanity (your own and others) 193
5.6 How to kill millions without noticing 205
5.7 Face-to-face experiences 209
5.8 Drawing more general lessons 221

CONCLUSION 223

BIBLIOGRAPHY 233

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Acknowledgments

I argue here that it can be rational for a person to trust another's advice about what to do. I am convinced of this in part because of the sagacious advice I received from my committee of advisors. No doubt this work would have be much weaker without their generous assistance; no doubt this work would have been much stronger had I more closely followed their advice.

First I want to thank Candace Vogler, who helped me from beginning to end. In many ways she has been the ideal advisor: her suggestions and recommendations always were offered in a helpful spirit; her patience withstood several drafts that give new meaning to the word 'rough'; the directions in which she has guided me were appropriate.

It was during Daniel Brudney's course on the philosophy of law that I first started thinking about why it might make sense for an agent to defer to another person's view about what to do. His encouragement and support convinced me both to pursue this question and to think more about why it is important. Michael Forster's valuable and insightful criticisms helped me to acknowledge the limited nature of the conclusions to which my arguments entitled me. And Martha Nussbaum was most

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

generous and helpful for making numerous suggestions that markedly improved the final product.

I also want to thank the participants in the Ethics Workshop and the Contemporary Philosophy Workshop at the University of Chicago. I greatly appreciate their constructive criticism, and I learned many things about how to do philosophy from their examples. I am unable to single out all of those who made significant contributions, but I want to thank Lauren Tillinghast, Erik Curiel, and Eric Brown for providing me with extensive written comments on drafts of multiple chapters. In many ways, my fellow students have been tougher on me than anyone else, and so my deepest debt is to them.

I gratefully acknowledge the financial assistance of the University of Chicago as well as the Mellon Foundation.

Besides acknowledging the professional and financial support I have received, I should also acknowledge the many important personal influences on my life these past few years. But I will do that in more appropriate fora.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

Although it is fair to say that my dissertation is a work in moral philosophy, I do not argue for a particular view about what sorts of things are good and bad, right and wrong. That is, I am not concerned here to argue whether things like pleasure, knowledge, or friendship are good.

Instead I am explicitly concerned with questions located in both metaethics and moral psychology, particularly with connections between the two. Metaethics is that area of philosophy concerned with questions about the status of ethics—whatever its content turns out to be—about its place in the world, about the meaning of our ethical discourse, and about our relation to it. Moral psychology is that area of philosophy lying at the intersection between moral philosophy and the philosophy of mind, concerned with those features of the mind involved in leading a moral life.

It is my contention that our metaethics will suffer if we fail to have an adequate moral psychology: our philosophical thought about morality needs to rely upon a realistic picture of moral perception, judgment, deliberation, and action.[1] Otherwise, our thoughts about the nature and place of morality are likely to exhibit

[1] Cf. "Modern Moral Philosophy" in Anscombe 1981.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

unimaginativeness, caricature, and a preoccupation with theoretical elegance. In other words, our metaethics won't be about *our* ethics, and this of course is a defect.

In Chapter One, I focus upon two related pieces of moral psychology. The first concerns the fact that the typical deliberator, when faced with a situation that gives her some pause, often consults others for advice. She does not decide what to do simply by thinking about what would be best; rather, she wants to know what others think as well. Others' advice is important to her, not only because without it she might overlook some consideration she thinks relevant, but also because she thinks that others may have a better view about what to do than she herself has.

When an advisee appropriately trusts her advisor's advice, I argue that the advisor *knows* what the advisee is to do. I justify this description by drawing comparisons between the trust an advisee places in an advisor, and the trust an inquirer places in an informant. Recent work in epistemology has convincingly displayed that we have the concept of knowledge in large part because we need to identify those well-positioned to tell us the facts.[2] Those who know whether *p* tend to be those from whom inquirers about whether *p* can learn the truth. Similarly, I argue that those who know whether to ϕ tend to be those from whom those deliberating about whether to ϕ can get sound advice. Well-positioned advisors can know what one should do.

[2] Craig 1990 and Williams 1972.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But this naturally leads us to think about what in fact characterizes a well-positioned advisor. Why should a person ever defer to someone else's thoughts about what to do? The deliberator would need to be able to tell herself some story about why the advisor is better placed to figure out what to do than she herself is. Either the advisor has some quality, or the deliberator lacks some quality, useful for figuring out what to do. I consider various candidates for what such a quality could be, and I argue that often it makes sense for a deliberator to defer to an advisor, when the former recognizes that the latter has undergone some *life-experience* that is likely to improve one's practical judgment. Thinking about the reasons we trust the advice of others, then, leads us to the second piece of moral psychology I focus upon.

By thinking about whose advice we should trust, we can see that the rationality of our actions is informed and guided not only by what we want, and by the information we possess, but also by the various life-experiences we undergo. The things we have done, the events we have lived through, the circumstances we have found ourselves in, the pleasure and the fear we have felt as we navigate through the unusual practical problems life throws into our laps—*these* things shape our present deliberations, and they do so in ways that can make the resulting actions more reasonable. Our life-experiences can truly be educative.

Unfortunately, these two features of deliberation do not receive due attention in contemporary metaethics, especially in discussions about whether morality is objective. One who dives into this literature is quickly surrounded by the increasingly

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

subtle distinctions philosophers make in order to stake out new positions: realism, antirealism, quietism, irrealism, error theories, projectivism, quasi-realism, secondary-quality realism, theories which combine realism with relativism, those that embrace objectivism but reject realism, even those which distinguish subjectivism from anti-objectivism, thereby making it possible to be both a subjectivist and an objectivist. One wonders whether the multiplicity of available options enables the philosopher to embrace exactly the right position, or whether it just makes it more difficult to see the forest for the trees. It strikes me that the explosion of answers to the question about whether morality is objective has not been matched by a explosion of resources brought to bear upon the question: the variety of argument output has not been matched by a variety of argument input, and so the responsibility for innovation lies with the cleverness rather than the resourcefulness of philosophers.

So in Chapter Two, I begin to take up the age-old question about whether morality is objective. I first argue for a particular way to understand the claim that morality is objective. I defend the common view that moral truths are objective only if they have explanatory power, only if the best explanation for why we think, for example, that some act is wrong actually makes reference to the fact that that act is wrong. But it is difficult to determine whether morality passes this test, and so I argue that the best way to tell whether moral truths explain our moral views is to see whether *im*moral views can be explained in a way that shows them to be an error or mistake. The defender of the objectivity of morality needs to show that at least some

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the parties to moral disagreement fail to have cognitive contact with the supposedly objective moral truths. He needs to show that the immoral are incorrect (which is not quite the same as showing that the immoral are immoral). Otherwise, the moral objectivist's principal claim is—as Bernard Williams claims—mere bluff.

One question that naturally arises here concerns the weight of the explanatory burden now upon the shoulders of the moral objectivist. Does he need to be able to convince the immoral person *herself* that she is in a poor position to grasp truths about how to live? Or need he show only more generally that the immoral are in error?

In Chapter Three, I argue that the moral objectivist need not convince the immoral person herself that she is in a poor position to grasp practical truths. I show that a person can fail to recognize her own deliberative deficiencies. I do this in the context of arguing that there can be external reasons for action: an agent can have a reason for action that does not serve her antecedently existing desires. Part of my argument for this conclusion establishes not only that a person may fail to recognize the reasons she has; she also may fail to recognize that she does not possess the qualities helpful for recognizing her reasons. The moral objectivist, then, need not bring the immoral person to see that she is not well-positioned to grasp practical truths. Rather, the moral objectivist need only display that the immoral person lacks qualities helpful for thinking well about how to live. And this can be done even if the immoral person disagrees with the cogency of this display.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In Chapter Four, I consider various specific ways a moral objectivist might actually exhibit the cognitive deficiencies of the immoral. Unfortunately, the deliberator one finds in the literature typically deliberates simply by bringing her *own* views to bear on her practical problem. Other available resources tend not to be dismissed so much as overlooked. And the paucity of materials the moral objectivist has in his toolbox leaves him unequipped to complete his task.

For instance, sometimes we can show that a person's practical views are incorrect because she lacks information necessary for making a good decision. At other times, we can show that a person's practical views are incorrect because she holds background philosophical theories that are unsound, theories that distort one's thoughts about one's proper activity. But the proportion of cases in which either of these kinds of explanation is available is probably small; in many if not most cases, we cannot chalk up another's immorality to poor information or unsound philosophical theory.

If this were the end of the story, the case for the objectivity of morality would look rather weak. Why should we think that moral truths are objective if we cannot even begin to explain why so many people (including at times ourselves) fail to apprehend them? What gets in the way? Better to think, it seems, that morality does not possess the objective status possessed by other domains, domains for which we can see how a theory of error might go.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But if we keep in mind the enriched picture of rational deliberation I develop in Chapter One, we find that we have more resources at our disposal for showing that immorality is a product of error. Recalling that we tend to trust the advice of those who have undergone significant life-experiences, we should be led to think that the lack of life-experiences can impair our practical judgment. Sometimes we can show that someone's thoughts about what to do are in error because they are inexperienced, sheltered, or otherwise naïve. Indeed, the advisee who trusts another's advice often thinks this way about *himself*—he justifies trusting the advice of his advisor because he believes that her experiences have improved her views about how to live, and that he, by contrast, is somewhat inexperienced about the whole matter.

A similar kind of explanation may be available, then, when trying to account for why the immoral have false views about how to live. It is tempting to think that such views often result from the lack of certain kinds of educative life-experiences. Sometimes people fail to have proper practical views, not because they lack information, nor because the philosophical theories they hold are unsound, but because their views are informed by too thin a range of life-experience.

In Chapter Five I adopt a strategy different from but complementary to the strategy deployed in the preceding chapter. Instead of linking the lack of experiences to the inability to grasp truths about how to live, I explore linking the lack of experiences to the inability to put one's views about how to live into practice. The idea here is that we often fail to live as we think we should, and part of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

explanation for why this is so includes the fact that we often fail to have the sorts of experiences that enable us to live up to our ideals. Sometimes we err because we lack integrity, and sometimes the lack of certain kinds of experiences fosters and sustains these forms of psychic division.

Specifically, I discuss the ways in which having face-to-face encounters with those whom one's actions affect enables one to act as (one thinks) one should. I focus upon how various psychological, technological, and institutional factors enabled physicians and other supporters of Nazi Germany to set aside or repress any humane thoughts they happen to have, factors that interfered with opportunities for them to see the faces of those whom they wronged. The lack of these kinds of encounters made it easier for them to wall off any correct moral views they had. And so they became wrongdoers in large part because their selves were fractured, because they lacked integrity. By displaying how the lack of one particular kind of life-experience can impair one's deliberations, I hope to make it more plausible that many if not most instances of immoral action can be traced to the lack of educative life-experiences.

Of course, there is still plenty of room for skepticism about the moral objectivist's claim. I simply hope to make a skeptical stance seem a little less tempting. It is exceedingly difficult to assess how much immorality can be explained as an instance of error. In fact, I argue that we lack good grounds for drawing any robust conclusions about whether morality is objective. For the objectivity of morality turns out to be partly a quasi-empirical matter, depending upon whether

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

those who underwent a full range of educative life-experiences would put correct moral views into practice. And this is something we are hardly in a position to know.

But we may be able to secure the thesis that *some* kinds of immorality can be explained as instances of error. Perhaps we can show that some kinds of immorality, such as violent injustice, are instances of error, even if we are unable to say anything conclusive about other kinds. And so one possibility that emerges—one that I broach in the Conclusion—is that some parts of morality are objective, even if other parts are not.

Now a caveat. I accept some popular distinctions in the course of my arguments, distinctions one probably wants to question. The idea that there is a line separating moral from nonmoral concepts, facts, views, and so on, is a simplification, perhaps a harmful one. Anyone who wants to argue that morality lacks the objectivity possessed by science, mathematics, or whatever first has to show or presume that morality is different from these things, that such lines can be drawn. I suspect one could attack such an argument by showing that to draw such lines is to misunderstand morality. But here I pursue a different route, a more internal critique of the argument against moral objectivity.

Finally, I should say something about the place of the present work in the contemporary philosophical literature. Although the germ of this work—the connections among advice, life-experience, and moral objectivity—has been developing for some five years now, the form of the final product owes much to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

work of Bernard Williams. As I have tried to articulate my arguments, I have found that Williams has also addressed these topics in essays some of which only recently have become widely available.[3] And so my arguments are largely concerned with the extent to which I think Williams is wrong. But debate with Williams is fruitful only because we already agree about many of the larger issues that surround the present debate. We find the same topics interesting largely because we share similar opinions about many other matters. And so the express disagreement rests upon a deeper philosophical agreement.

---

[3] Bernard Williams 1995a.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter One

Advice and action

1.1 Rational agency

How does a rational person decide what to do? Philosophical pictures of rational agency vary tremendously, but this should be no surprise. For a topic so frequently covered, we should expect nothing less than a diversity of views about what being a rational agent amounts to. Perhaps second only to the role of inquirer, philosophical writing portrays the human being as an agent, as one who does things.

A cursory survey of leading philosophical positions reveals that the rational agent, in deciding what to do, consults her own "desires"[1], "second-order desires"[2], "subjective motivations" or "ground projects"[3], "strong evaluations"[4], "principles"[5], "ideals" or "self-understanding"[6], "conception of how to live"[7], "considered

[1] Davidson 1980.

[2] Frankfurt 1988.

[3] Williams 1981.

[4] Taylor 1985.

[5] Harman 1977.

[6] Anderson 1993.

[7] McDowell 1979.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

judgments"[8], and/or "practical identity"[9]. The details of the picture vary depending upon the particular philosopher in question, as just about any issue of a journal of contemporary philosophy will make clear. The question whether our reasons for action spring from the orectic or the cognitive faculties of the mind is all the rage. But these squabbles belie a deeply shared assumption: in all of these pictures the agent decides what to do by getting her actions to line up with her *own* psychological states. Whether by design or by default, deliberation is portrayed as essentially monological; the contemporary philosopher's agent operates in the same milieu as the Cartesian doubter.

But I want to argue that this is at best a half-truth. The rational person does not consult merely her own thoughts in deciding what to do. She also consults the thoughts of others. Further, sometimes she trusts another person's advice, even when her advisor hasn't brought her to see for herself that what he says to do is best.

Now it might be thought the rational person would precisely *not* act in ways she didn't see for herself as best. It might be thought a hallmark of rationality to act only in ways that one can fully endorse. It might be thought that an agent violates her autonomy in putting the decision about what she is to do in the hands of another.

---

[8] Rawls 1971.

[9] Korsgaard 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I will argue, however, that these thoughts are misleading, and that the person who trusts the advice of another person may be more reasonable than the person who goes it alone.

1.2 TESTIMONY

Fortunately, our task is eased by the fact that philosophers have recently rectified some of the more formal distortions associated with earlier pictures of rational *inquiry*, distortions that I believe also plague contemporary pictures of rational agency. For too long, philosophy had neglected or de-emphasized the ways in which an individual's beliefs depend upon what others tell her. Instead, it had been thought that one should believe only those things that one can find out oneself, without relying upon the word of others.[10] Understanding why it can make sense to trust another person's testimony as to how things are, I believe, will help us see why it also can make sense to trust another's advice about what to do.

The critique of individualist epistemology I want to focus upon is Professor Craig's recent monograph.[11] Instead of entering the time-honored debate about how to reply to the skeptic, or the recent protracted debates about the definition of

[10] See, for example, Descartes [1637] 1985.

[11] Craig 1990.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

knowledge, Craig instead steps back to ask what the point of the concept of knowledge is. Why do we *care* whether someone—let's call her Agnes—knows *p*?

We might care whether Agnes knows *p* because *we* do not know whether *p* or not-*p*, and we want to find someone who does.[12] Does Agnes know whether *p*? If so, she would be a good person to ask if we ourselves want to find out whether *p*. People who know can inform those who do not. This suggests we have the concept of knowledge in large part because of our frequent need to find informants about various matters. Except for those in the business of educating others—and this may explain how the distortion came about—we are inquirers much more often than we are examiners.

Craig develops this link between the concept of knowledge and the need for informants in many interesting ways. He begins by considering an "ordinary situation", an uncontroversial example that displays some important features of our concept of knowledge:

> Human beings need true beliefs about their environment, beliefs that can serve to guide their actions to a successful outcome. That being so, they need sources of information that will lead them to believe truths. They have 'on-board' sources, eyes and ears, powers of reasoning, which give them a primary stock of beliefs. It will be highly advantageous to them if they can also tap the primary stocks of their fellows—the tiger that Fred can see and I can't may be after me and not Fred—that is to say, if they act as informants for each other. On any issue, some informants will be better that others, more likely to supply a true belief. (Fred, who is up a tree, is more likely to tell me the truth as to the whereabouts of the tiger than Mabel, who is in the cave.) So any community may be presumed to have an interest in evaluating

[12] Williams 1972.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> sources of information; and in connection with that interest certain concepts will be in use. The hypothesis I wish to try out is that the concept of knowledge is one of them. To put it briefly and roughly, the concept of knowledge is used to flag approved sources of information.[13]

Our sources of information, the things which typically provide us with true beliefs, include not only our various cognitive faculties, but the cognitive faculties of others. Often we don't have a belief about whether *p*, cannot (or would rather not) find out for ourselves, and thus need to rely on informants to find out whether *p*. So the inquirer wants an informant such that: either *p* and the informant believes that *p*, or not-*p* and she believes that not-*p*.[14]

But this is not all the inquirer wants. He wants to find an informant who not only believes *p* when *p*, and believes not-*p* when not-*p*, but someone for whom *he* will believe that *p* if she tells him that *p*, and for whom he will believe not-*p* if she tells him that not-*p*. The inquirer wants to find someone who satisfies some condition which correlates well—as the inquirer believes—with telling others whether *p*.[15] Craig calls such a property an *indicator property*. Sometimes an indicator property will be very simple. In the case mentioned in the quote above, the indicator property is the fact that Fred is up a tall tree. I am more likely to believe what Fred tells me about whether *p* if I think that he has the right indicator properties for correctly

---

[13] Craig 1990, 11.

[14] Craig 1990, 12.

[15] Craig 1990, 13.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

believing whether *p*. Knowers of *p*, then, will tend to have some property detectable to persons to whom it is not detectable whether *p*, a property that will mark them as likely knowers of *p*. (Not that all knowers of *p* will have in common any one property, but that for any *p*, some properties will correlate more strongly with true-believers of it than with others.) That this is so follows from approaching the concept of knowledge from the point of view of the inquirer.

Thus "a very large part of the art of acquiring correct information consists in being able to recognize the sort of person . . . that will have the right answer."[16] The inquirer asks himself "who knows whether *p*?", and looks for people who have some property X, a property that he thinks is likely to be possessed by someone who will tell him the truth as to whether *p*. It does the inquirer no good to know that property X correlates with telling the truth about whether *p*, if he cannot determine who has property X. Likewise, it is for naught if the inquirer can tell who does and does not have the property X, if he does not also realize that having such a property is correlated with telling him whether *p*; that is, if he does not realize that property X is an *indicator* property with respect to whether *p*. Thus a good informant is not simply a person who has a true belief about whether *p*, but a person who also can be identified by inquirers as such. Such a person is the archetypal knower. The knower of *p* is the person from whom others can find out that *p*.

---

[16] Craig 1990, 26.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.3 THE POSSIBILITY OF ADVICE

Do philosophical pictures of agency display a similar defect? Do they one-sidedly focus upon the ways in which a deliberator depends upon her own views about what to do? Do they ignore the ways in which a deliberator listens to what others say to her?

As the catalog I sketched above indicates, the standard pictures of rational deliberation tend to limit their resources to the psychological states *of the deliberator herself*. The portrayed agent consults her own desires, her own principles, her own conception of how to live, her own practical identity, . . . , and acts in the light of these. The fact that other people have different views seems to matter only insofar as the agent's action affects other people. So if Agnes already has as one of her principles, "I shouldn't hinder others from ϕ-ing", and she learns that someone else wants to ϕ, this piece of information about another's psychological life might make it unreasonable for her to do certain things. But if *someone else* thinks "Agnes shouldn't hinder others from ϕ-ing", the reasonableness of the various courses of action open to Agnes seems not to have changed. What other agents think seems to be irrelevant.

But here philosophy is out of step with practice: when a person is in a quandary about what to do, she often turns not inward but outward, consulting those

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

close to her for advice. In deliberating she can recognize the limits of her own views, leading her to seek advice from others. Just as an inquirer might defer to another's judgment about how things are, so too might a deliberator defer to another's advice about what to do.

Many questions may be asked about the relationship between a deliberator and an advisor. Is this really a common feature of practice? Is it ever rational for a deliberator to defer to the advice of an advisor? Does the deliberator really defer to the *advice* of the advisor, or does he defer to something else the advisor tells him? Does the deliberator who trusts the advice of the advisor think that the advisor knows what he should do? And, finally, what indicator properties does a deliberator look for in an advisor?

Let us begin with a rather simple scenario, one where it seems incorrect to ignore the benefits of advice. We reasonably trust another person's advice in scenarios where the advice is essentially about *how* to do something. We can often identify people who can inform us about how to do things we don't know how to do on our own. For instance, an electrician knows how to rewire kitchens. And we are able to identify competent electricians even if we ourselves are unable to rewire kitchens. If I want to find out how to rewire my kitchen, I can ask an electrician for instructions. She might break the problem down into smaller parts, each of which I can do. By doing each of these smaller parts, I will thereby rewire my kitchen. It can be quite reasonable for me to trust what the electrician says about what to do, granted

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that one of my ends is to rewire the kitchen. I can't do what I need to do simply by thinking about the problem all by myself. Rather, I need some help, and I can get it by trusting the electrician's advice.[17]

So it seems that the electrician has knowledge about what I should do. She will tell me to do A, B and C, and so I will learn how to rewire my kitchen. But perhaps it will be doubted whether her knowledge is anything more than a particular form of ordinary propositional knowledge. Perhaps the technical advisor knows only particular and general facts which the advisee doesn't, facts the knowledge of which enables one to ϕ. And so the advisor can either give the advisee the information he lacks, or, rather and more simply, tell the advisee how to ϕ. The idea here is the advisee could ϕ himself if he knew these facts known by the advisor.

But I suspect that this objection would not bear analysis. That is, I suspect that we can show that the person who knows how to do something knows more than just information which is often too cumbersome to share.

However I will not undertake to argue this here, for even if we show that knowing-how-to-ϕ cannot be boiled down to knowing-that-*p*, there is still the more relevant question whether knowing-how-to-ϕ is the only form of knowledge potential advisees seek to tap.[18] Do advisors seek advice only about technical matters? Do

---

[17] See Craig 1990, sec. XVI.

[18] Ryle (1949) forcefully and to my mind convincingly argues that knowing-how-to-ϕ cannot be boiled down to knowing-that-*p*.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

deliberators want to find out from others only the best means to their already-adopted end?

Suppose I already know *how* to rewire the kitchen, but I still am wondering *whether* to rewire the kitchen. Doing so would be rather expensive, and I'd really like to use the money instead to visit my parents. But the risk of electrical fire is greater if I leave things as they are. I know that I ought to rewire the kitchen in order to reduce the risk of fire, and I know that I ought to leave things as they are in order to visit my parents, but I still don't know which to do. It seems that it would be good if I *could* find someone who knew which I should do. People are often in a quandary as to what to do, where this does not boil down to a quandary about how to do something. And if there were those who could inform deliberators as to what to do, the deliberator would do well to heed such advice.

But it seems that there are several problems with thinking that this kind of advice is possible. First, it is difficult to swallow the notion of this kind of informant, of someone who has knowledge about whether one should ϕ. It is commonplace to think that there aren't experts about this kind of thing.[19] In particular, practical thought doesn't seem like science or mathematics or hierarchical religion, fields in which there are knowledgeable people whose pronouncements we can trust. Practical

[19] Eggerman 1979. See also Bernard Williams, "Who needs ethical knowledge?" in Williams 1995a, 205.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

thought seems to be more egalitarian than that.[20] (You don't need to make a man a philosopher in order to make him a man.)[21]

Second, there is something suspect about ϕ-ing simply because somebody tells you that you should ϕ. It seems that rational action is always autonomous action: one needs to understand and endorse the grounds of one's action if it is to count as truly reasonable. We can legitimately trust the journalist who tells us that the Supreme Court has ruled that abortion is legal; it seems, however, we cannot legitimately trust the person who tells us that abortion can (not) be a reasonable thing for someone to do. Making ϕ-ing your end simply because someone else told you to ϕ, while it may wind up preventing you from acting poorly, doesn't seem very rational. This is not to say that there is no role for talking things over with other people. Often we benefit from listening to what others have to say, seeing whether they can *convince* us that we should do one thing rather than another.[22] But this is not analogous to Craig's characterization of knowledge: no one reasonably *defers* to another's judgment about what to do.

A third objection is argued forcefully by Bernard Williams:

---

[20] Cf. Coady 1992, 71.

[21] See Coady 1992, 75: ". . . . I, for one, would no sooner think of consulting your average moral philosopher over a genuine moral problem than of consulting a philosopher of perception about an eye complaint."

[22] Cf. Coady 1992, 71: "When someone tells me that what I am proposing to do is immoral, I do not react by asking for his credentials but for his reasons."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> [A] good advisor of one person need not be a good advisor of another—and not merely in the sense that there are some people who cannot be advised by anyone. An advisor, and the person seeking advice, may not share the same presuppositions. Someone could be a capable and insightful advisor, to Catholics, for instance, who accepted the value of chastity, but be no use to someone who did not; in the opposite direction (so to speak) a seeker after advice might think that some well-regarded and shrewd advisor displayed a louche and opportunistic consequentialist outlook. In these ways, ethical knowledge, to the extent that it is identified through the advisor model, remains local. Moreover, the advisor model itself cannot be extended to identify a kind of knowledge that could itself overcome these difficulties. You cannot identify an advisor by marks that do not already include the degree of ethically shared outlook that would enable the person seeking advice to trust and understand the potential advisor. . . . This means that the model does very little for the larger concerns of cognitivism. Cognitivism's question has often been expressed simply by asking whether there is any ethical knowledge or not, but in fact it has typically been concerned with the hopes of resolving the kinds of disagreement that separate from one another the local practices of advice under shared ethical presuppositions.[23]

If advisor and advisee must accept the same considerations as relevant—if they must largely share the same perspective—then the resources that are available to the advisee seem to be limited indeed. For if the insight that the advisor can offer involves the advisee's grasping that certain courses of action fall under various "thick ethical concepts" (e.g. that a course of action would be cowardly, or kind, or a betrayal), then this casts doubt on whether it is proper to say that the advisor has *knowledge* about what the advisee is to do. An advisor can help one see that a certain action falls under a thick concept only if that concept belongs to one's culture, in Williams' example, the culture of Catholicism and its like-minded neighbors.

[23] Williams 1995a, 207-8. Beware of Williams' use of the word "ethical". By that he does not mean "moral", as his example of the shrewd, opportunistic advisor illustrates. (Recall the consigliere!) I think it is closer in meaning to the idea of "practical".

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Advisors from other cultures apparently do me little good, and this seems to be for two reasons.

The first might be that I wouldn't be able to grasp the content of their advice, couldn't be brought to understand the meaning of the thick concept under which the proposed action falls.[24] But the second and more important reason I will not benefit from such an advisor is that I can identify a trustworthy advisor only if she and I *already* significantly share a practical outlook. I will not trust an advisor who deliberates using different thick concepts, for she will not have most of the indicator properties I look for.

Since any guidance the advisee receives will come through the vehicle of thick concepts, we cannot be sure that the advisor can really inform the advisee (or anyone) about what to do. The problem seems to be that there is no way to adjudicate fairly between deliberating with the thick concepts of one culture, such as one's own, and deliberating with the thick concepts of another culture. In practice, of course, one will think about what to do, and seek advice about what to do, with those of one's own culture. Indeed, this is part of Williams' very point. But the fact that any truly potential advisor will deliberate largely with the same thick concepts as the advisee seems to mean that the reflective advisee cannot get help as to whether his thick

[24] One could certainly disagree with Williams' conclusion by attacking this argument. I pursue a different strategy.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

concepts are a better way of thinking about what to do than those of some other culture.

So if I need to figure out whether ϕ-ing would be chaste, I can ask someone who should know, someone who also deliberates using the concept of chastity. And such a person may very well know whether ϕ-ing would be chaste. But if I want to know e.g. whether chastity is a virtue, whether thinking about various courses of action in terms of chastity is a reasonable way to deliberate, whether I should do the chaste thing, then Williams thinks that an advisor will be of little help, for any advisor I am willing to trust will not tell me anything I don't already think. Because our sources for advice are in this sense local, the model of an informant seems structurally unable to provide a vantage point from which to determine whether one should deliberate in the light of one set of thick concepts rather than some other set of thick concepts. After all, no one is omnicultural.

These three related objections suggest that Craig's conception of knowledge cannot be straightforwardly applied to practical thought. The individualistic nature of deliberation seems to make practical thought unlike other kinds of thought.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.4 Advice and Knowing What to Do

But I think we can overcome these problems, and show that there is a legitimate role for a practical advisor, one that conforms to Craig's account of knowledge. I will take up each of the above objections in turn.

First, even if there are no theoretical experts on how to live, there still may be trustworthy advisors, as long as a trustworthy advisor's knowledge need not be of the theoretical variety. The indicator properties for appropriate advisors may be very different from those for other kinds of informants. More on this shortly.

Second is the concern whether one's ϕ-ing can really be rational if one ϕs simply because one trusts someone else's advice. Now this objection needs to be addressed carefully, for there is some truth to it, but it will take some time to extract it. The first thing that I think should be said is this: if the advisee ϕs because the advisor tells him that ϕ-ing is the thing to do, then, while it perhaps would have been better had he arrived at his decision to ϕ all on his own, it is probably better that he ϕ than that he ignore the advice. Of course, this presumes that the advisor is actually right about whether the advisee should ϕ. But if this presumption is correct, then we should be able to see that the puzzled inquirer would be *more* responsible to heed the advice of the advisor than to refuse to take advantage of the resources available to him.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Furthermore, it is still up to the advisee whether to accept the advice proffered. Things would be different if one were *coerced* into ϕ-ing. But when one freely chooses to trust another's advice, one thereby exercises one's autonomy. To see this, reflect upon whether an *inquirer* necessarily forfeits his intellectual autonomy when he defers to the judgment of someone who possesses (what he takes to be) appropriate indicator properties. It is very plausible to conclude instead that an inquirer sometimes exercises his intellectual autonomy when he decides whom to trust: he might believe an independent journalist's report of the recent coup, as opposed to the report put out by the Party, for (he thinks) independent journalists tend to possess more qualities associated with speaking the truth about such things. It can be perfectly reasonable for the inquirer to trust the journalist. Similarly, a deliberator's trust in another's advice might be perfectly reasonable if that person possesses appropriate indicator properties. And so the trusting of another's advice need not violate one's own autonomy.

Nevertheless, there is still something very peculiar about (freely) choosing to ϕ simply because an admittedly excellent, trustworthy advisor recommends it. Even if she is correct about whether he should ϕ, it seems that he has to see *for himself* that ϕ-ing is appropriate in order for him to reasonably do what she says. The faith the advisee places in the advisor cannot be blind, not like the faith I possess when e.g. I trust the FDA that these pills are not toxic. An agent's need to understand the reasons

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

for his actions seems to undermine the possibility of his reasonably trusting advice about whether to ϕ.

In order to fully answer this objection, we will need to draw some distinctions. These distinctions tend to simplify, perhaps oversimplify, the nature of and the circumstances surrounding advice. But I hope that these simplifications help more than harm our thinking about the topic.

Suppose I am wondering whether to ϕ, and am looking for some help on the matter. Someone I'm inclined to trust tells me that I shouldn't ϕ because *p*. (For purposes of illustration, suppose ϕ-ing is the action of 'eating this pot of stew' and *p* is the proposition that 'this pot of stew is not kosher'.) Now, after hearing what she has to say, I still do not see for myself that I should not ϕ. Our views about the propriety of ϕ-ing may differ for several reasons.

One reason is that I may have my doubts about *p*. She says I shouldn't ϕ because *p*, but is *p* really true? Here the advisor and I are in dispute about what the facts are. I have trouble seeing why I should not ϕ because I do not accept her view of the circumstances. I might choose to defer to her view about whether *p*, but this would not be to trust the *advice* she offers. Rather, it would be trust her expertise in some other area, for instance, her knowledge of Jewish dietary laws and/or of the ingredients and preparation of this pot of stew in front of me.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Another reason I might fail to see for myself that I shouldn't ϕ is that, while I grant that p is true, I think that it is *irrelevant* to my decision. I might recognize that p is true, but I don't think it has any bearing on whether I should ϕ. Or, more interestingly, I recognize that p is true, but whereas my advisor says that p is a reason not to ϕ, I think that p is a reason to ϕ. (Perhaps I want to rebel against my Orthodox parents.) The advisor and I here seem to have different sets of values, and it indeed would be quite odd for me nonetheless to trust what she says about the impropriety of ϕ-ing. It appears that in order for me to reasonably trust her advice, I have to see that the considerations that she thinks support her recommendation indeed *are* considerations that so support her recommendation. And I also have to see for myself that the reasons she cites are *relevant* to my decision in roughly the way she explains them to be. So far, our objector is on firm ground.

The third reason I might fail to see for myself that I shouldn't ϕ is that, while I grant that p is true, and that it is relevant in the way my advisor says that it is, I don't fully see why she thinks it is so *important*. For example, while I grant that p is a reason for me not to ϕ, there is also q (e.g. my military officer has ordered me to eat the food given to me), which seems to me to be a more important reason. My advisor agrees with me that q is true, and that q is relevant in the way that I think it is. But she maintains that the fact that p is more important than the fact that q, and so that I shouldn't ϕ. I am confused about what to do, but if forced to make a decision on my

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

own, I would probably conclude that *q* is more important than *p*, and therefore that I should ϕ. But I may nonetheless decide that the advisor is better than I at determining the relative importance of *p* and *q* with respect to my ϕ-ing. I may decide to refrain from ϕ-ing because I trust her judgment.

I suspect that this can be a reasonable thing for one to do. Even if one cannot see the relative importance of some consideration that one finds to be both true and relevant, sometimes it can make sense to trust the advice of someone who *does* think that such a consideration is fairly important. Of course, this all depends on whether we can identify anyone who is qualified to dispense advice about what is relatively important. But assuming that we can, it might make sense—it would not necessarily be irrational—to heed her recommendations about what to do.

So the second objection has application when the disagreement between advisor and advisee centers on issues of truth or relevance. But when the advisor and the advisee disagree only about what is relatively important, the second objection may not necessarily succeed—it seems possible that an advisee can reasonably defer to an advisor's advice when they differ only over which of several relevant considerations is most weighty.

Now we must take up the third objection, the one that arises from the long passage I quoted from Williams. Williams argues that the advisee can seek advice only from those who inhabit the same practical (he says "ethical") world, from those

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

who deliberate with the same batch of thick concepts. An advisee can be led to see or trust that ϕ-ing would be cowardly only because the concept of cowardliness is a concept he already uses in deliberation. This implies, Williams argues, that the advisor can inform the advisee only about whether these thick concepts apply. Specifically, the advisor does not demonstrate that she can actually inform the advisee to do or to avoid doing the cowardly thing.

Now a proper rejoinder to the problem that Williams presents must acknowledge the partial truth of his analysis. Perhaps it is true that one

> cannot identify an advisor by marks that do not already include the degree of ethically shared outlook that would enable the person seeking advice to trust and understand the potential advisor.

We arrived at a similar conclusion in our response to the second objection.

But there are some differences between that conclusion and the analysis Williams offers, differences that prove significant. For we argued that it could make sense to trust another's advice about what to do only when the only differences between advisor and advisee concerned the relative importance of considerations they both acknowledge as true and relevant in the same way. So, yes, they must in some sense share a practical outlook. But the kind of knowledge that the advisee credits to the advisor is not knowledge involving whether a particular thick concept applies. In fact, the advisee has to see *for himself* the truth and the relevance of the considerations the advisor sees. Otherwise, the advisee will look elsewhere for help.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Instead, the rational advisee needs help in deciding what to do in light of these agreed-upon considerations, and it is here that he is open to trusting the advice of someone else.

So, admittedly, any advisor he might trust will not differ radically from him in practical outlook. But this means that the advisee asks himself, "Who knows whether I should ϕ?", and that he looks for someone who probably has this knowledge. If the advisee is going to trust the advisor's advice about anything, it will be about whether to ϕ. So the fact that the sources of advice must be somewhat local yields the slightly paradoxical result that the rational advisee will defer only to the advisor's judgment about what he should do, and not to her judgment about what thick concepts apply to proposed courses of action.

It appears, then, that Williams has confused the indicator properties possessed by the advisor with the content of the advice given. The *indicator properties* of the good advisor will necessarily include (but will definitely not be restricted to!) her beliefs about the facts and their relevance, the content of which should roughly match those of the advisee. Otherwise, Williams' negative argument gets a grip. So, yes, if the advisee thinks that there is something to be said for chastity, then the advisee will want only an advisor who thinks this too (presuming that this bears on the particular deliberative problem the advisee is facing.) But the content of the advice given is not the same as the content of the beliefs that advisor and advisee share. How *could* it be,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

given that the advisee needs to align his actions with the advice of the advisor? Rather, the advisee wants to know what to *do*, and this is what the advisor may be in a position to tell him. So the advisor might tell the advisee that the fact that ϕ-ing is chaste is less (or more) important than some other relevant consideration. And so the advisee might trust what the advisor says about what to do.

Thus it is correct to say that the advisee thinks the advisor has knowledge about what to do, at least if we understand the concept of knowledge in the way that Craig recommends.

But does the advisor *in fact* have such knowledge? Is it really reasonable to defer to anyone's advice about what to do? Heretofore I have been concerned with the *possibility* of rationally trusting advice, with whether it even made sense to do what a presumably qualified advisor recommended. Now it is time to examine the plausibility of that presumption, to see whether there are those whose advice it would make sense to heed. We need to articulate what qualities could make someone an appropriate advisor, one whose advice merits deference. If we are successful, we will show that it can be reasonable for a deliberator's actions to be regulated by *another person's* views, that standard pictures of deliberation are indeed incomplete.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1.5 Some Dead Ends

When one is thinking about whether to ϕ, what qualities *does* one tend to look for in a potential advisor, if any? Are there indicator properties for practical knowledge? We have already specified that the advisor and advisee must share a view of the facts and their relevance, if the advisee is to trust her judgment about what to do. But that is hardly sufficient to be a good advisor; otherwise, the advisee could 'advise' himself better than anyone else could.[25] What *else* is important?

First, an advisor should be someone we trust; that is, we not only trust that she has sound judgment, but we also trust that she does not have any ulterior purposes. She will not tailor her advice in order to benefit herself, or others toward whom she is partial. She has our interests at heart.

Second, an advisor may really be just a motivator. Often we know what to do, but, for various reasons, we do not act. We need someone to remind us why we should do it, remove paralyzing doubts, assure us that we are on the right track. The advisor tells us what we really knew all along, but puts us in a better position to execute such knowledge.

Third, an advisor may be able to help us see what the facts we already know imply, what conclusions we may reasonably draw from the facts we know. Perhaps

[25] Michael Smith has suggested that "suitably idealized, we are in fact the best people to give ourselves advice." The question, then, is what constitutes the forms of idealizations. See Smith 1994, 151.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

we are wrought up, too emotional to think straight. We need someone to keep our thinking on the rails, to prevent us from under- or overreacting, to keep us from committing fallacies.

Fourth, an advisor is often imaginative. She is able to think of considerations we might overlook, of how various possible courses of action might affect other people. She may even introduce possible courses of action we had not yet contemplated. Discussions with such an advisor are likely to be fruitful, for we will end up seeing things we had not seen before, enabling us to make better decisions.[26]

All of these ways an advisor can help one decide what to do, however, are ways that the advisee himself can check up on. In all of these cases, it seems the advisee can evaluate the quality of the advice for himself: the advisee doesn't really need to *trust* the advisor. For instance, the advisor might point out that the advisee is committing the *post hoc ergo propter hoc* fallacy, but then the advisee will see for himself whether she is right. He might have failed to catch the fallacy without the advisor's help, but he doesn't need to rely on the advisor's word that it really is a fallacy. Or, an advisor may highlight a possible course of action he had been overlooking, but then he can consider for himself whether it really is an available

[26] Eggerman (1979) offers a longer list: a good advisor is sincere and conscientious, knowledgeable about the facts of the case, disinterested, non-dogmatic, proficient in detecting fallacious reasoning, proficient in tracing out probable consequences, does not make problems seem simpler or more complex than they really are. My comments about my shorter list apply equally to his.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

option. The advisor's imaginativeness is something he can mine—he uses it to make up for his own lack of imagination—but once extracted it is his own gem.

In none of these cases, then, do we find a situation where the advisee does what the advisor advises because he, at least in part, trusts her practical judgment. Although he is aided by the advisor, he can re-create monologically whatever it is the advisor tells him, evaluating for himself how it is to influence what he does. The relationship between advisor and advisee here is not the same as the relationship between Craig's truth-seeker and informant. The truth-seeker identifies the informant by her indicator properties, and then trusts what she says regarding the matter at hand. The logical/imaginative advisor points out when the advisee is overlooking something—either an error he has made or an option he has not considered—and the advisee sees the thing he was overlooking, now better equipped to deliberate on his own. In these cases, the advisee need not *trust* the advisor's judgment about what to do.

(This highlights some of the differences between the aspects of practical thought I am exploring and the criticism of monological practical reasoning forwarded by Habermas. In his dialogical ethics, Habermas argues that the validity of norms depends upon whether all affected could accept the consequences and side-effects of their general observance.[27] Thus we actually need to see what norms would

[27] Habermas. "Discourse Ethics: Notes on a Program of Philosophical Justification" in Habermas 1990.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

emerge from a dialogue of affected parties; we cannot monologically determine which norms are in fact valid. However, affected parties other than myself apparently do not inform me about anything that I should trust. Specifically, they do not tell/inform/advise me what to do; I do not defer to their practical judgment. I trust them only about whether they accept proposed norms. So I *do* need to trust fellow interlocutors when they tell me about their own psychological states; dialogue informs me about this piece of information. But for Habermas, the process of dialogue doesn't really inform me about what to do; it just provides me with information that I combine with a practical norm I supposedly already accept.)

1.6 Life-Experience: Doing Well

Nevertheless, I want to maintain that some advisors do have qualities signaling that their advice is trustworthy. To do this, I want to consider and develop a few remarks Aristotle makes about advice in his moral philosophy. In the *Ethics*, he quotes Hesiod approvingly:

> Far best is he who knows all things himself;
> Good, he that hearkens when men counsel right;
> But he who neither knows, nor lays to heart
> Another's wisdom, is a useless wight.[28]

But how do we identify those who counsel right? Aristotle answers:

---

[28] Aristotle, *Nichomachean Ethics*, trans. Ross and Urmson, 1095b10-13. Quote from Hesiod's *Works and Days* 293-7.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> Therefore we ought to attend to the undemonstrated sayings and opinions of experienced and older people or of people of practical wisdom not less than to demonstrations; for because experience has given them an eye they see aright (*Ethics*, 1144a11-14).

Aristotle tells us that we should listen to what experienced people say about what to do, for their experiences have enabled them to see the truth about the things they tell us. He suggests that their experiences have educated them in ways that can neither be demonstrated nor fully explained. Advisees are to use the fact that such people are experienced as an indicator property for knowledge about what to do, as something that is strongly positively correlated with practical wisdom.

Now one might think that this remark of Aristotle's is difficult for us to take very seriously. The respect naturally accorded to the aged in Aristotle's day no longer survives for many moderns. Nevertheless I think Aristotle points toward something that is of use to us. In order to see that this is so, we will need to examine Aristotle's own reasons for thinking that experience improves practical judgment, as well as other reasons to accept this conclusion. While practical knowledge may not increase proportionally with age, we do tend to seek advice about what to do from those—young or old—whose life-course and experiences shed light on *our* situation. Aristotle seems to acknowledge a similar qualification himself:

> A young man is inexperienced in the actions that occur in life . . . it makes no difference whether he is young in years or youthful in character; the defect does not depend on time . . . (*Ethics*, 1095a2-7)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

When in a dither about what to do, when we are looking for an informant about how to proceed, one of the many characteristics we look for in an advisor is experience about the actions that occur in life. We want someone who has been through what we are going through, or considering going through. They may have learned something from which we can benefit. The experiences of others can serve as a guide in our own deliberation.

Now Aristotle thinks the kind of experience had by the trustworthy advisor is the experience of having performed particularly *virtuous* actions. On his view, one needs certain virtues in order to possess the kind of practical judgment had by the good advisor, and one can acquire these virtues only by performing virtuous deeds. So deliberators can identify good advisors in part by noting whether they have performed many virtuous actions.

Aristotle forges the link between the practice of good actions and sound practical judgment by connecting each of them with the notion of pleasure:

> For an activity is intensified by its proper pleasure, since each class of things is better judged of and brought to precision by those who engage in the activity with pleasure; e.g. it is those who enjoy geometrical thinking that become geometers and grasp the various propositions better, and, similarly, those who are fond of music or of building, and so on, make progress in their proper function by enjoying it . . . (*Ethics*, 1175a30-36)

In order to acquire the capacity to act properly, one must take pleasure in so acting. Taking pleasure in a sphere of activity can enable one to see what is of value in it. Pleasure has a cognitive aspect to it, making one aware of the worth of the objects of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

one's passions (*Ethics*, 1105b21ff., *Rhetoric* 1370a27-28). It draws one's attention toward those passions whose objects are particularly attractive, those which seem to constitute good reasons for engaging in that activity. And the pleasure arising from seeing the point of an activity furthers one's sense of what doing that activity well consists in. Similarly, taking pain in performing appropriate actions destroys one's ability to determine which actions *are* in fact appropriate (*Ethics*, 1140b11-16, 1144a35-36, 1175b14ff.).

So by repeatedly acting in ways that others tell him are just, a person can come to understand how the control he exhibits for the sake of others is appropriate, and to take pleasure is so doing. Practicing good action "has cognitive powers, in that it is the way we learn what is noble and just."[29] The agent comes to see the point of just actions, the way in which acting such to render each person her due is a noble end. Once he gets some sense of this, he begins to take pleasure in acting in appropriate ways, providing both the motivation and the cognitive capacity for sustaining a life-long pattern of just activity.[30] Further just activity brings about a fuller understanding of its importance, which further enhances the pleasure taken in acting virtuously. Practice, understanding, and pleasure grow simultaneously, each reinforcing the others. Eventually, this pattern of activity can become second nature,

---

[29] Burnyeat 1980, 72.

[30] Aristotle, *Nichomachean Ethics*, II.3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

at which point it can be said that one has acquired the relevant virtue: it has become a state of one's soul.[31]

I must confess that I find Aristotle's specific account of how virtuous action improves one's practical judgment somewhat puzzling. Does the wise person fighting for justice really take pleasure in doing what he does? We must presume that by pleasure Aristotle does not mean self-satisfaction, the pat on the back that one gives oneself for resisting temptation. The ethical person, on the other hand, will generally empathize with those whom his virtuous action benefits, and so can take pleasure in *their* improved situation, divorced from any thought of his own *noblesse oblige*. But empathy does not seem to play a prominent role in Aristotle's ethical philosophy, and while importing it helps to make sense of what Aristotle says, it also smacks of anachronism.

But we are not doing the history of philosophy now; if the person who is experienced at performing virtuous action can feel how others are feeling, then this is relevant to our present problem. Whether Aristotle has this in mind is unimportant. The person who has experience performing virtuous actions may have a better sense about how actions affect others, and this is something that cannot easily be represented propositionally, cannot be cashed out in terms of true sentences. At best, the experienced person may be able to say only things like "φ-ing would not simply

[31] Aristotle, *Nichomachean Ethics*, II.5.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

please her, it would please her *very much*." Perhaps only a person with a strong sense of empathy could fully understand the content of what is meant by such an utterance. Others would have to trust her when she said that this consideration was the most important one. It seems that an advisee can identify a good advisor, then, by noting that she has performed many virtuous actions, for there seems to be a link between this and grasping how one's actions affect other people.

But the chain composed of the link between performing virtuous action and a heightened sense of empathy, and the link between a heightened sense of empathy and improved judgment about what to do—*this* chain does not seem very strong. There are probably many counterexamples to this pattern, many people who perform virtuous actions without gaining a strong sense of empathy, as well as many empathetic people who lack sound practical judgment. It's a good start, but we should continue to search elsewhere for additional reasons why those with various kinds of experiences are likely to dispense sound advice.

1.7 Life-experience: not doing so well

Curiously, perhaps we look for advisors whose experience consists not in having performed obviously virtuous actions, as Aristotle maintains, but in having performed actions whose appropriateness is controversial. If we are wondering whether we should ϕ, we may be likely to trust the advice of someone who has

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

already ϕ-ed. Perhaps her ϕ-ing taught her something about whether ϕ-ing is something worth doing. Those inexperienced with ϕ-ing may be able to benefit from what she has to say.

I believe that this is in fact a very common reason why we tend to trust the advice of others. Let me offer an example.[32] In the late 1940s, a relatively large proportion of American veterans coming home from the War became very active in the U.S. political process. Many of them ran for Congress in 1948, and as a result, there was an unusually high turnover in its membership the following January. One hundred eighteen new members, many of them War veterans, were eager to improve the country's situation. Specifically, they felt that fighting against the Axis gave them "a special sensibility" about when it was appropriate to send American troops to battle, and when it was not. The motto of this cohort was "They sent. We went.", as though they were in a better position than their predecessors to assess the appropriateness of going to war. Though proud of their own service, they felt that they would be more reluctant to commit soldiers to dangerous situations than their predecessors had been. Having experienced first-hand the horrors and the sorrows of war, they felt they were better judges about the costs of war, costs whose greatness they wouldn't have been in as good a position to assess had they themselves not

[32] The following information is from an NPR broadcast January 8, 1993. The reporter was interviewing Charles Bennett, a Florida Congressman who came to Congress in January 1949, and was the last of his cohort to retire.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fought. Their experiences gave them a better sense, they felt, of how important it is to avoid these atrocities.

War is an activity whose propriety is often controversial. Although there are many wars that are obviously a bad idea, this is not always so. Now those who must decide whether to go to war should be, among other things, well-informed and imaginative. But we also want them to be experienced; naïveté about these things should be avoided. Those whose familiarity with war originates only from history books are more likely to fail to comprehend its toll.

There are a couple of dangers this strategy faces, however. First, it is fairly common for those who do questionable things to rationalize their actions. They wish to have a positive self-image, and so they are more likely to think that what they did was appropriate. Those who war, then, may think that warring is appropriate more often than they should. Second, war can be *so* dreadful, that the judgment of those who must experience it becomes impaired. Mental illness and military service aren't uncommon bedfellows. And, indeed, it is common to think that unwise action generally corrupts one judgment.

But if one is not impaired by the experience, it is plausible that one would be a better judge of the dreadfulness of war if one had experienced it for oneself. One might come to know just how hellish war can be, something one wouldn't fully grasp if one instead were, say, a war historian. This improvement in practical judgment would not arise solely from additional knowledge about the facts of war, but also

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

from a stronger sense of the *cost* of the losses as expressed by whatever the facts happened to be.

This suggests that if we are deliberating about whether to respond to some crisis with military force, consulting the views of those who have themselves fought is likely to be appropriate. Of course, the decision whether to go to war depends on so many considerations, not the least of which is the difficulty of gathering relevant information, that no one consideration is likely to outweigh all the others. And we may be confident enough of our own views such that we feel secure in rejecting the opinion of the war veteran. But all I hope to have shown is that (and why) he may be in a better position to judge properly, that we can envision how one's view about the importance of such things is likely to be improved through such experiences. Even though the war veteran's experience stems from actions of questionable appropriateness, we can see that his judgment about whether we in our present circumstances should fight can seem authoritative. He may have knowledge about this matter.[33]

Something similar may be said for those who have undergone suffering or other misfortunes. Often one who has survived tragedy commands the respect of others. The thought is that their difficulties have given them a better sense of what is

[33] We will return to the case of the experiences of the war veteran in Chapter 4, where we will be concerned with the value of their autobiographical writings.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

important in life or what really matters. Their improved understanding is something we can rely upon in our own deliberations.

For example, consider what one American newspaper printed about the writings of Primo Levi:

> A triumph over the experience of Auschwitz and over our reluctance to read what might give us pain. Levi has given us the sense of what it was to survive, not as victims, but 'as men made to follow after excellence and knowledge.' Failure to read such a work is to consign a river of human experience to oblivion.[34]

Another reviewer writes:

> Primo Levi is that rare individual, a survivor who can write about his experiences yet keep a sense of balance and proportion . . . and by listening to him, so will we all learn.[35]

The sentiment expressed by both of these reviewers is that Primo Levi's life-experiences taught him about how one is to live, and that his undergoing hell-on-earth can teach *us* how to live better as well. Ignoring what we can learn from his life would be to "consign a river of human experience to oblivion."

Much the same can be said for other survivors of the Holocaust, as well as for many others who have suffered great evil. Those of us fortunate enough not to have witnessed the evil that men do frequently acknowledge that those who have possess a seriousness that is worthy of respect, and that they have something to say. Indeed

---

[34] From the *Chicago Tribune*, reprinted inside the front cover of Levi 1993 [1958].

[35] From the *San Francisco Chronicle*, reprinted inside the front cover of Levi 1993 [1958].

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

their words are even taken to be authoritative. Having seen for themselves the darkness within humanity, they are less prone to romanticizing things, they have a better grasp on reality, on what is important.

There are dangers here, of course. Levi writes of how the Nazis not only corrupted themselves, but also their victims. It would be a mistake to think of the suffering as simply refining the character of those who survived. Stories of concentration camp prisoners acting like amoral animals abound; indeed, to animalize men was a goal of the camps. Both perpetrators and victims of evil are degraded in various ways. These considerations place limits on what we can say about the relationship between suffering and practical judgment. Sometimes suffering corrupts practical judgment, sometimes it improves it, and sometimes suffering does both. This makes it more difficult to identify which sufferers are trustworthy.

More difficult, but not impossible. We can be confident that a particular person has not been corrupted by the evil they have suffered; it shows up in how they carry themselves, how they talk about their past. It would be wrong to think that we do not have much to learn from such individuals. There may be special problems associated with our seeking advice from those who have survived some great evil. Elie Wiesel writes, "Ask any survivor, he will tell you, he who has not lived the event will never know it. And he who went through it will not relate it, not really, not entirely. Between his memory and his reflection there is a wall - and it cannot be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pierced." True enough, but, as Wiesel himself demonstrates, the rest of us can often benefit from whatever the sufferer *can* tell us about his reflections.

Once our attention is drawn to the fact that we can and perhaps do place special stock in the words of those who have suffered the horrors of totalitarianism, we will soon recognize that we also regard those who have endured more mundane hardships as nevertheless having something to say. All sorts of people learn various things from their difficulties, and these lessons are not always ones that must remain private.

Not only can suffering and hardship improve one's practical judgment, the *threat* of suffering can as well. Consider a different example. It can be convincingly argued that people can agree on all the pertinent facts about the fetus, her mother, and their circumstances, and yet still reasonably disagree about the propriety of abortion.[36] Although for some people one's view about the propriety of abortion seems to follow quite readily from the information one believes one holds, for many people this is not so. For those of us who fall into this latter category, can we identify certain types of experiences which might disclose the relative importance of some consideration we agree is relevant, in order to consult those who have undergone such an experience about the propriety of abortion?

---

[36] Wertheimer 1971.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It seems plausible that we could. Those who have had such importance-disclosing experiences might include: a woman who has had an abortion in order to be able to pursue other options; a similarly-motivated woman who sought an abortion but was unable to procure one; her child who now realizes the circumstances of her birth; a man who wanted the mother of his child to have an abortion, but she didn't (and he is told by the President that he "must take responsibility for the children he has brought into this world"[37]); his child; a man who didn't want the mother of his child to have an abortion, but she did; and so on. As a generalization, those who go through these crises probably have a better sense of the relative value of the considerations at stake than those of us who have not. Having, say, their personal autonomy on the line, some of them can testify as to the importance or unimportance of having the ability to control one's own body. Alternatively, those who know that their birth took place only in virtue of their mother's inability to procure an abortion might have a better sense of how important or unimportant it is that abortion results in the death of a human embryo/fetus.[38] As a result, it seems appropriate that we give their attitudes about the importance of these considerations some weight when thinking about the propriety of abortion. We may be moved to alter our notions about what is most important in the light of their views, views we may come to trust. We

---

[37] *State of the Union Address*, February 17, 1993.

[38] The Reverend Jesse Jackson has wrestled his whole life with this knowledge about the circumstances of his birth. See Frady 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

may be moved to act differently in light of their testimony. Those of us whose lives and liberties haven't been so threatened would probably do well to listen to those who have.

Action whose appropriateness is controversial, as well as the undergoing of hardship and suffering, appear to be types of experiences that can foster practical judgment. The advisor whose authority stems from the suffering or hardship she has undergone, or from her participation in actions whose propriety is unclear, is not quite like Aristotle's *phronimos*, the person whose practical wisdom arises from and is reinforced by the pleasure he takes in virtuous activity. She is more like a food tester, or a canary in a coal mine, one who has emerged from harm or potential harm having learned lessons she can share with others. This analogy is not perfect, but it vividly depicts the differences among the various credentials an advisor might be thought to have. People learn from these incidents, and others can often benefit by heeding what such people have to say.[39] So it would seem that Aristotle was on the right track when he recommended trusting the advice of the experienced. Further reflection reveals that there are a *variety* of types of experiences that shape one's views about what to do. It still seems plausible to say that we would do well to heed the "undemonstrated sayings" of those whom we recognize as having undergone

[39] Levi, for one, is convinced that in fact "*no* human experience is without meaning or unworthy of analysis."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

experiences that are likely to have improved their thoughts about how to live. We just need to widen our views about the kinds of experiences that are relevant.

1.8 Conclusion

Earlier (1.4), I claimed that sometimes advisees think that the advice of others is worth trusting. More recently (1.6, 1.7), I have suggested that, under certain conditions, this can be reasonable to do when the advisors have appropriate life-experiences. Now if Craig's explication the concept of knowledge is largely correct (1.2), then we may conclude that sometimes advisors *do indeed* have knowledge about what to do.

Aristotle's remarks help bring to light something we knew all along, but that we may forget when doing philosophy. In deliberation, we often turn to others. Sometimes we trust what they say. Sometimes we do this because they are more experienced than we are; their experiences are one of several indicator properties that mark them as good advisors. And sometimes it actually is reasonable for us to trust them on these grounds.

We now can see why standard pictures of rational deliberation are incomplete insofar as they depict the rationality of an action to depend only upon the agent's own views. The advisee may still believe exactly what she believed before she took the advisor's advice, want exactly what she wanted before she took the advisor's advice,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conceive of herself exactly as he conceived of herself before she took the advisor's advice . . . . The advisor may have convinced the advisee only to *do* something that she otherwise would not have done. And if this can be a reasonable thing for the advisee to do, then we can see how the reasonableness of the resulting action can flow not merely from the rationality of the advisee's own views—for *they* can remain the same throughout—but also from the advice of the advisor.

Thus neither an individualist epistemology nor an individualist moral psychology accurately portrays the ways in which our beliefs and actions depend upon what other people tell us. Just as it is often reasonable for an inquirer to trust another's testimony about how things are—reasonable even though the inquirer has not necessarily come to believe anything *else*—so too can it be reasonable for a deliberator to trust another's view about what to do.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER TWO

Understanding moral objectivity

2.1 METAETHICS AND PHILOSOPHICAL PSYCHOLOGY

In Chapter One, I argued that the rational agent sometimes defers to the advice of others. The importance of the differences between this enriched picture of rational agency and other, flatter pictures is not restricted to philosophical psychology. In particular, discussions about the nature of morality are affected by the picture of rational agency they assume. A Kantian understanding of agency naturally fits and supports a Kantian understanding of the nature of morality. Similarly, there are connections between hedonic theories of what motivates agents and utilitarianism. So too are there connections between a picture of rational agency that makes room for advice and certain philosophical thoughts about the nature and place of morality. In particular, I will argue that the fact that we are agents who take advice from the experienced is relevant to debates about the objectivity of morality.[1]

[1] Many Anglo-American philosophers now debate whether moral *realism* is true. It is unclear whether this shift in language has signified a change in topic. (But see Norman 1997.) I will usually speak of moral objectivity, but sometimes will speak of moral reality when discussing the views of others.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.2 Two thoughts about objectivity

More immediately, however. I will articulate why the objectivity of morality depends upon the proper understanding of ourselves. The fundamental claim made by the moral objectivist is that morality is, in some sense that needs to be articulated, part of the world, rather than an invention, construction, projection or expression of the mind. But these are formulaic ways of putting it, and it is likely that any successful argument for or against moral objectivity will also end up clarifying in what moral objectivity consists.

Nevertheless, some general remarks about the notion of moral objectivity can be made in advance of such arguments, remarks determinate enough to provoke some controversy. The objectivity of morality is sometimes characterized by the combination of two thoughts.[2] The first is that the content or the substance of morality—e.g. that marital fidelity is right, that animal torture is wrong—is generally *independent* of the shape of the moral views we happen to have. Morality and our conception of morality are not necessarily identical; our conception of morality is, at best, only a map of a domain independent of us. In this way, moral objectivism is a modest position for us to hold.

[2] See Wright 1992 and Railton 1984, both of which characterize moral realism in this way independently (apparently) of one another. This is especially interesting since Wright and Railton disagree about whether morality *is* objective.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But there is also a presumptuous element of moral objectivism, and this is that we are capable of getting things right, that our thoughts about morality can be correct, that we are such that morality can interact with us, and that it can give us *feedback* on how things stand in the moral realm. Even though our thought about morality is only a map, it can be an accurate map, and not accidentally so, but because morality somehow exerts an influence on our thoughts and lives.

These two thoughts are often taken to characterize the content of moral objectivity, thoughts which in turn can be understood in multiple ways. How should we understand independence? How could morality exert influence on us? The most promising strategy for understanding these thoughts, I suspect, is to understand how they operate and how they relate to one another in the context of the actual arguments produced by both objectivist and anti-objectivist. Toward this end I shall turn to an examination of some specific arguments against moral objectivity.

But first I want to consider one reason for querying whether we really need both thoughts in order to characterize the notion of moral objectivity. One might expect that the notion of moral objectivity is sufficiently characterized by the thought about independence alone, such that any concern about interaction is merely a distraction, a distraction that the opponent of moral objectivity will likely exploit in order to promote illegitimate doubt. If morality is independent of our thoughts, then perhaps this is enough to conclude that an objective understanding of morality is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

appropriate. Focusing on whether and how morality determines our thoughts seems, though interesting in its own right, beside the point.

There is some plausibility to this objection. Though the notion of objectivity is both obscure and fairly technical, mind-independence appears to characterize very closely what we typically mean by it. Yet debates about the mind-independence of morality do not proceed in isolation from debates about relations of dependence that *do* exist between mind and world. In principle, perhaps, our objector is correct: objectivity and mind-independence come to the same thing. But *we* only come to know whether morality is mind-independent by considering how it could nevertheless make some difference with us. If a mind-independent morality could not reveal itself to us, we would have no grounds for thinking that morality is independent of our thoughts. This approach, of course, rules out the possibility that there are moral truths about which we cannot know, perhaps because morality is too deep and profound for such wretched creatures as ourselves ever to understand. This approach rests on the assumption that if morality is mind-independent, we will still be able to discover its features. If this assumption is false, we will wrongly conclude that morality is not mind-independent, and thus wrongly conclude that morality is not objective. Nevertheless, we should not find this assumption unreasonable. While there may be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conceptual room for the position that we cannot discover what in fact is morally right, we ought not place too much hope in that situation actually obtaining.[3]

However, a closer examination of the relation between the thought about mind-independence and the thought about feedback reveals that the latter rests upon an assumption significantly stronger than the one just discussed. To accept the thought about feedback is to presume that moral objectivism is false so long as we cannot deliver a satisfactory account of how a mind-independent morality makes a difference in our thoughts and our lives. But of course, it is possible that we interact with morality without realizing just how this happens. Perhaps there are features of our lives which *are* best explained by noting what is good and bad, right and wrong, without our being able to articulate how this explanation goes. It could even be that we falsely believe that we do not so interact with morality. Thus if we interact with morality, there is no guarantee that we will be in a position to explain how this works. This approach rests on the assumption that if morality is mind-independent, we will be able to explain how we interact with it. If this assumption is false, we may wrongly conclude that morality is not mind-independent, and thus that moral objectivism is mistaken.

Nevertheless, I am willing to grant this assumption. While there is room for the position that we interact with morality despite not being able to understand how

[3] Whether there is conceptual room for this position depends on whether realism in the sense articulated by Michael Dummett is correct. See his "Realism" in Dummett 1978.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

this is so, it does not seem granting the anti-objectivist this much leeway insures his victory. Let us be charitable with this assumption, and see where he can go with it.

But we should articulate what it is that presses us into making the above assumptions. The worry motivating these assumptions is that we will believe that morality is mind-independent, and that we will make room for morality in our conception of the world, when in fact morality is only a construction of the mind. The burden of proof is thus on the objectivist to show that morality is independent of our conception of it.

Allocating the burden in this way is sometimes justified on the grounds of ontological parsimony. If we do not need to posit a mind-independent realm of moral facts in order to account for features of our moral life, then it is extravagant to do so anyway. Making room in our view of the world for moral facts that make no difference in our lives is to overpopulate the world with facts not needed. And for many moral anti-objectivists, overpopulation is one of the worst sins:

> If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe . . . How much simpler and more comprehensible the situation would be if we could replace the moral quality with some sort of subjective response which could be causally related to the detection of the natural features on which the supposed quality is said to be consequential.[4]

> If we can explain why it should come about that we have the A-concepts (e.g. moral concepts), with their horizontal connections, loosely fitting the B-descriptions (e.g. nonmoral descriptions), and in this explanation rely only upon our exposure to a B-describable world, then at least the metaphysical

[4] Mackie 1977, I.9. See his similar treatment of secondary qualities in Mackie 1976, 17-18.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> and epistemological motivations will be answered. We will be able to do away with a distinct area of A-facts, which troubled the metaphysician.[5]

> In explaining why we make the normative judgments we do, I found normative facts superfluous. . . . I do deny that there are normative facts . . . [6]

> Convenience to one side, we can explain just as well confining ourselves to the non-moral terms of the descriptions. And if we can, we should, in the interests of ontological economy.[7]

Any "entities or qualities or relations" whose existence we countenance must do some work; idle, useless entities, qualities and relations are to be cut out with Ockham's mighty razor. If we can make sense of what is uncontroversially so—e.g. that we have such-and-such views, that we act in such-and-such ways—without recourse to speaking of moral facts or entities or descriptions, then ontological parsimony requires that we exclude morality from our conception of what there really is.

I suspect that one who places much weight on the value of ontological parsimony may do so because of a fear of what is chaotic or unknown, or, alternatively, for *aesthetic* reasons. Nevertheless, I shall not mount any attack on the role this value plays in the allocation of the burden of proof between the objectivist and the anti-objectivist.

---

[5] Blackburn 1984, 162.

[6] Gibbard 1990, 122.

[7] Miller 1985, 527.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

An alternative justification for the approach I've been articulating rests not on any Quinean worries about overpopulating our ontology, but more directly on matters associated with the point about interaction. Here the concern is instead drawn from Wittgenstein, that we are in the grip of a picture that morality *must be* mind-independent if it plays the role in our lives that we think it does, but that in fact a proper understanding of our lives might reveal that this picture is not necessary after all. Perhaps it is a fantasy to think that we need the moral objectivist's particular conception of 'morality'—one in which it is mind-independent—in order to make sense of the role morality actually plays in human life. It would be easier to rid ourselves of this fantasy if we had a better grasp on why such a notion is an idle wheel spinning free of the mechanism that is human life, if we knew the way that things actually work, if we were *truly* realistic.[8] Hence the concern with the proper account of the relation between morality and mind. This is another way to understand the demand that we include the presumptuous thought as well as the modest one when we characterize the content of moral objectivism. If we see that the morality does not play the role we were tempted to think that it does in the presumptuous thought, our reasons for holding on to the modest one dry up.

So moral objectivity consists not simply in the modest thought about independence, but in the presumptuous one about interaction as well. Defending the

---

[8] Diamond 1991, 45.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

objectivity of morality requires that we have some account of how a mind-independent morality could nevertheless influence aspects of our lives.

2.3 Explaining moral views

Gilbert Harman has produced an argument against moral objectivity that isolates the problem faced by anyone hoping to show that a mind-independent morality aids us in understanding the nature of our moral views.[9] Harman asks us to consider the various "observations" we make. Observations, as Harman uses the term, are the particular opinions one forms in response to what one perceives; one does not consciously infer an observation from anything else. It is important to note that observations can be as theory-laden as you like; it is a virtue of Harman's argument that it does not rely on a dubious theory/observation distinction. Another virtue of his account is that an observation is not supposed to be the infallible Given upon which everything else is built. If Nancy the astrologer looks up into the sky, sees the planet Jupiter in a particular constellation, and immediately thinks to herself "I'm going to die tomorrow!", we shall rightly say that Nancy observed that she will die tomorrow, even if we think that what she observes is false.[10]

---

[9] Harman 1977.

[10] Thus it makes sense to speak of moral observations even if there is no special faculty for detecting moral truths.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Now Harman is concerned with how to determine the relationship between mind-independent facts and observations. He asks whether we need to make any assumptions about moral facts in order to explain the *moral* observations we make, observations about what is right and wrong, good and bad. He notes that our moral observations can either agree or conflict with the more general moral principles we hold. This is no different from the relation between scientific observations and scientific theory. But, as Harman notes, there remains the following disanalogy.

In science, we "need to make assumptions about certain physical facts to explain the occurrence of the observations that support a scientific theory."[11] A physicist may see a vapor trail in a cloud chamber, and think to herself: "There goes a proton." This thought could occur to her without any conscious reasoning having taken place; this thought could count as an observation in the relevant sense. Now this observation could help to confirm her theory about what a proton is like. Its status as confirmatory rests on inferring an explanation for why she observed what she did. The best explanation for her observing a proton includes not only the assumption that she was in a certain psychological "set" at the time of the experiment—one based on her beliefs about the scientific theory and the experimental apparatus—but also the assumption that there really was a proton going through the cloud chamber, which produced the vapor trail, which she observed as a proton. If

[11] Harman 1977, 6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

her having that observation could be equally well explained without any assumption about a proton, then her observation would not have been evidence for the existence of the proton, and thus would not have been evidence for the scientific theory. Her observation confirms the theory only because we cannot easily explain it solely on the basis of her psychological set at the time of the experiment. It is reasonable to assume, over and above her psychological set, that there really was a proton in the cloud chamber, which produced the vapor trail, which she saw as a proton.[12]

Harman contrasts this case with one involving observations about what is right and wrong. Suppose as you begin to walk down an alley, you suddenly see a bunch of kids set a cat on fire. You may think to yourself, "What those kids are doing is wrong." This counts as your making a moral observation. Now Harman asks whether the wrongness of their activity explains your moral observation in the same way that the presence of the proton explains the physicist's observation. Are moral observations explained by facts about rightness and wrongness in the same way scientific observations are explained by facts about subatomic particles? One way to explain your observation that the kids are doing something wrong is that the kids are *in fact* doing something wrong, in the same way that the best explanation of your observing that there is a cat being burned is that a cat is *in fact* being burned, and that

[12] Thus Harman is a scientific realist. The contrast which Harman makes between ethics and science could be attacked by those who have a different view of science. I, of course, do not pursue this route.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the best explanation of the physicist's observing that there is a proton in the cloud chamber includes the fact that there is a proton in the cloud chamber.

But Harman maintains that explanations that refer to moral facts are unsatisfactory. For your moral observation can be explained more plausibly and more simply in the following way: you already hold certain moral principles which imply that burning animals for kicks is wrong, and these principles in combination with your observation that the kids are burning the cat for kicks explain why you observe that their activity is wrong. We needn't assume that the hoodlums' action actually was wrong in order to account for your indignation. The most reasonable explanation of your observation is in terms of your moral principles, principles which you probably acquired as a result of your upbringing and socialization. You have been brought up to associate wrongness with some *non*evaluative property you observe, and your moral observation can be adequately explained by making reference to both this nonevaluative property and this association.

Now Harman thinks that an observation can provide evidence for a theory because the truth of that observation can be part of an explanation of why that observation was made. When a theory-laden observation is explained by the facts as described by the content of that observation—when my theory-laden observation that *p* is explained, in part, by the fact that *p*—then the observation confirms that theory.

Since the best explanation of the physicist's observation that there is a proton in the cloud chamber includes the fact that there is a proton in the cloud chamber, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

physicist's observation counts as evidence for her scientific theory. Since we needn't assume any moral facts in order to explain why you made that moral observation, your moral observation is not evidence for your moral theory. Your observation that the kids are doing something wrong is not evidence that they are doing something wrong. Since we presumably have no other grounds for thinking that morality delivers feedback (recall the second thesis I referred to earlier), then we shouldn't think that there is any mind-independent fact stated by the sentence "What those kids are doing is wrong." We have no account of how a mind-independent morality could make a difference in our lives. And so, Harman concludes, there are apparently no moral facts.

Harman's argument is swift, ingenious, and to the point. It remains to be seen whether it is valid or sound. There are many questions about the soundness of Harman's argument which could be discussed, but I first simply want to investigate only the validity of his argument. If the best explanation for why we make the moral observations we do need not refer to moral facts, should we conclude that there are no objective moral facts? For ease of expression, I will call this test for objectivity *the reference-in-explanation test.* We need to see whether the reference-in-explanation test is an appropriate articulation of the feedback requirement. For now I will grant Harman the truth of his premise, that the best explanation for why we have the moral observations that we do does not refer to any moral facts.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But perhaps I should say a word about the plausibility of his premise in order to motivate my discussion of the argument's validity. Harman claims that theoretical physical entities are needed to explain the physicist's observation, but moral qualities are not needed to explain your moral observation. Obviously the truth of these claims depends heavily on what alternative potential explanations are available. Specifically, there must be some explanation that does not refer to moral qualities for why we make moral observations superior to the explanation put forward by the moral objectivist, viz. that we observe that what the kids are doing is wrong simply because what the kids are doing is wrong. Harman has never produced a detailed rival explanation (although other anti-objectivists have), confining himself to broad, sweeping claims about the transmission of moral principles through education, training, and upbringing. But we can see how the details of an alternative explanation might go, and, unless we are already wedded to an objectivist picture of morality, this alternative explanation will likely have some plausibility. It is not obvious that such a rival explanation would be wrong.

The same cannot be said in the case of the physicist's observation. Few would argue that the best explanation for why the physicist observed a proton in the cloud chamber has nothing to do with whether there in fact was a proton in the cloud chamber. While doubt in the case of moral observation seems at least plausible, if not likely, doubt in the case of the physics observation is harder to muster up. Thus it seems worthwhile to investigate the validity of Harman's argument, in order to see

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

whether his claims about the explanatory inefficacy of moral facts, if true, should lead us to reject moral objectivism.

In the rest of this chapter, I shall examine a barrage of arguments attacking the validity of this argument. First, we will look at an argument that exploits the supervenience of the moral upon the nonmoral; then, at an argument that questions the relation between the explanatory efficacy of putatively moral facts and their normativity; then, at an argument that takes our judgments involving artifacts as a test case for the suitability of the reference-in-explanation test; and finally, a pair of arguments that question whether it is possible to explain at least some of our moral observations without being committed to their truth. I will argue that each of these arguments, despite having able proponents, fails to show that the reference-in-explanation test is an inappropriate gauge for determining whether morality is objective. I will argue that the link between the objectivity of morality and the explanatory power of morality has not been snapped.

2.4 Objectivity, Causation and Counterfactuals

So let us now examine the validity of Harman's argument, which maintains that since we do not need to refer to moral facts in order to explain our moral views, morality is not objective. We can question whether the lack of reference to moral facts—which for present purposes we are granting—is equivalent to the lack of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

dependence on moral facts. That is, even if the explanation of the moral observation in question does not explicitly refer to any moral facts, perhaps moral facts are still somehow relevant to the explanation, relevant in a way that supports moral objectivism.

One strategy is to argue that a gap between reference and relevance opens up because of the way in which the moral realm supervenes upon the nonmoral. The moral character of a situation is supervenient upon its nonmoral facts, if it is impossible for the moral character of that situation to change without some of the nonmoral facts changing as well. Or, alternatively put, if two situations have different moral characteristics, then they cannot be identical in all non-moral respects.

Susan Hurley economically articulates how the supervenience of the moral realm upon the nonmoral realm opens up the gap between relevance and reference.[13] She wants to show that our moral observations can still depend on the existence of moral facts, even if the best explanation for why we have those observations makes no reference to those facts. She wants to show that we can be said to have moral *knowledge*, even when our beliefs are not explained by moral facts. And this is supposed to show that even if our moral observations are not explained by moral

[13] Hurley 1989.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

facts, they still depend on moral facts in a way which undergirds the objectivity of morality.[14]

To this end, Hurley adopts a counterfactual account of knowledge.[15] On this account, Selena knows that *m*, if and only if

(1) *m* is true,
(2) Selena believes *m*,
(3) if *m*, Selena would believe *m*, and
(4) if ~*m*, Selena wouldn't believe *m*.

Selena knows that *m* if she truly believes that *m* ((1) and (2)), and if her belief that *m* covaries with the truth or falsity of *m* ((3) and (4)). The covariance of Selena's belief that *m* with the truth of *m* is one way to understand the requirement that the fit between mind and world not be accidental, that Selena's true belief constitutes knowledge rather than a lucky guess. It is, at least on the surface, a different way of understanding the feedback condition than is Harman's, who demands that the explanation of the belief refer to, rather than merely covary with, its content. But let us see whether these differences deliver different results.

Does this model of knowledge make room for moral knowledge? We are supposing that there are moral truths; thus it seems that, if *m* is a moral truth, (1) can

[14] Williams (1985) considers a view of moral knowledge similar to Hurley's view. But he carefully distinguishes moral knowledge from moral objectivity: he thinks that we can have moral knowledge even if morality is not objective.

[15] See Nozick 1981.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

be met.[16] Further, we are not here entertaining the possibility that there are moral truths which no one believes; we should accept (2) and (3).

(4) is the tricky one, for it is a live question whether we would have the moral beliefs that we do even if moral truths were otherwise. Our moral beliefs seem to depend upon biological, sociological and other cultural influences, not upon moral truths. And so if moral truths were otherwise, it appears that we would nevertheless maintain pretty much the same moral beliefs we now have. And this casts some doubt over whether (4) can be met.

To see whether (4) is true, then, Hurley suggests we let *m* stand for some moral judgment, such as "It is wrong for those kids to burn that cat!", and let *n* stand for the judgment expressing the conjunction of the relevant nonmoral facts concerning this situation. Let B*m* stand for Selena's belief that *m*. Let *c* stand for whatever counts as the best explanation of B*m*, where, let us assume to make Harman's premise true, $c \neq m$.

Then, Hurley marshals an argument based in part on the supervenience thesis discussed earlier,

(5) If ~*m*, then ~*n*,

as well as a premise stating that *c* is necessary for B*m*

---

[16] That is, we are assuming that ethical discourse meets all the requirements set forth by a minimalist theory of truth, for ethical sentences have appropriate syntactic features, and their use is constrained by appropriate standards of discipline. For more on the ways in which truth and objectivity are distinguished, see Wright 1992.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(7) If ~*c*, then ~B*m*.

Her argument will work if she can show both that

(6) If ~*n*, then ~*c*,

and that the resulting string of counterfactuals are transitive:

(5) If ~*m*, then ~*n*. (Supervenience thesis)
(6) If ~*n*, then ~*c*. (Remaining question)
(7) If ~*c*, then ~B*m*. (Best explanation of B*m* (*ex hypothesi*))
(4) If ~*m*, then ~B*m*.

If these counterfactuals are true and transitive—they are not *necessarily* transitive—then we may conclude that (if ~*m*, then ~B*m*), and the link between the belief in question and the truth of that belief is not snapped by the fact that moral properties are not referred to by *c*.

(5) and (7) are obviously met; (5) expresses the supervenience of the ethical upon the nonethical, (7) is true by stipulation. So the remaining questions are whether (6) is true, and whether these counterfactuals are transitive.

Consider the issue of transitivity first. Unfortunately, we have no guarantee that the counterfactuals are transitive: the state(s) of affairs referred to by the consequent of (5) may not be among the state(s) of affairs referred to by the antecedent of (6), even though they both are symbolized by "~*n*". Similarly, the state(s) of affairs referred to by the consequent of (6) may not be among the state(s) of affairs referred to by the antecedent of (7), even though they both are symbolized by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"~*c*". Interpreting the antecedent of a subjunctive conditional is far from straightforward.

To illustrate, if we want to evaluate whether the statement "If I weren't a philosopher, I would be a lawyer" is true, we need to think about what would be true if I weren't a philosopher. Now there are many conditions under which I would not be a philosopher: if I had been kicked out of graduate school; if I had been raised in a family with an established medical practice; if I had won the lottery: if I had been killed in an automobile accident as a child: if my parents had never met; if life never developed upon the earth. When we consider whether such a conditional is true, we ignore the kind of situations specified by these latter examples of possibilities. Even though I wouldn't have been a philosopher if life had never developed upon the earth, it still might be true that if I weren't a philosopher, I would be a lawyer. We ignore these more remote possibilities to focus upon more local possibilities, possibilities in which we change only a minimum amount of what is now true in order to render the antecedent of the conditional true.

The problem about transitivity arises, then, if we consider the following two sentences, both of which we may assume are true:

> If there had been a global nuclear war in the 1980s, I wouldn't be a philosopher.
>
> <u>If I weren't a philosopher, I would be a lawyer.</u>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Yet, it is false that if there had been a global nuclear war in the 1980s, I would be a lawyer. Transitivity fails because the states of affairs referred to by the consequent of the first conditional are different from the states of affairs referred to by the antecedent of the second conditional. Each clause points to different sets of possibilities.

So looking at the subjunctive conditionals Hurley speaks of, we need to think about the possibilities specified by *~m*, *~n*, and *~c*. Sticking to our earlier example, what are the possibilities specified by the antecedent "if it were not wrong for those kids to burn that cat"? The supervenience thesis guarantees that *n* would be false in all such possibilities. Let us suppose that the minimum we have to change about the actual world to make *m* false makes *n* false in the following way: the cat in question is very slowly bleeding to death and is obviously in great pain, the kids cannot take it to a veterinarian, and, seeking to put it out of its misery as quickly as possible, using the only means they can think of to this end, they ignite the cat. Then it wouldn't be wrong for those kids to burn the cat.

Now we are in a position to see whether (6) can be met. If the cat in question were slowly bleeding to death, etc., would whatever explains Selena's present belief that burning the cat is wrong still be true? Suppose Selena's belief that *m* is caused in part by her belief that the kids are burning the cat for fun. This nonethical belief interacts with Selena's other thoughts to produce Selena's belief that *m*. So (6) is not met if Selena would still believe that *m* if the kids were just trying to put the cat out of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

its misery. And so she would think that what the kids were doing was wrong even if it wasn't.

But surely it is possible for Selena's beliefs about the nonmoral facts to track the nonmoral facts. If the kids were trying to put the cat out of its misery, then Selena very well may no longer believe that what they were doing was morally wrong. Selena may find out the truth about what the kids were really up to. And if Selena doesn't believe that the kids were burning the cat for kicks, then she probably won't believe *m* either. Learning about the kids' intention, Selena may come to think that it wasn't wrong for them to burn the cat. She may no longer believe *m*.

Thus whether Selena knows that *m* boils down to whether Selena knows which nonmoral propositions of the situation the moral propositions supervene upon, knowing the truth-values of those nonmoral propositions, and having her moral beliefs sensitive to this latter bit of knowledge. Selena can have moral knowledge, even though *m* doesn't cause Selena's belief that *m*. And so on the counterfactual theory of knowledge, ethical knowledge is possible. Harman's argument appears invalid.

Another way to see the charm of this position is to examine a couple of contrasting examples offered by Nicholas Sturgeon.[17] Suppose Bobbi thinks that Pat's homosexuality is wrong. Also, suppose that if Bobbi weren't confused about her own

---

[17] Sturgeon 1992.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sexual identity, she wouldn't think that Pat's homosexuality was wrong. Now the features upon which the putative wrongness of Pat's homosexuality supervenes—the disposition to engage in certain sexual acts—do not vary with the explanation of Bobbi's views in the right way in order for Bobbi to know that Pat's homosexuality is wrong. That is, if Pat wasn't dispose to engage in those sexual acts, it would still be true that Bobbi remained confused about her own sexual identity.[18] So (6) does not hold, and thus Bobbi fails to have moral knowledge.

Contrast this with the case where Quincy thinks that slavery is wrong, and Quincy wouldn't think that slavery is wrong only if he didn't think that slaves were fully human, or that slavery makes the slaves miserable, or that it prevents them from realizing their capacities for self-development and self-respect, and so on. Now the features upon which the wrongness of slavery in fact supervenes (n)—the misery of the slaves, etc.—are *identical* with the content of the beliefs of Quincy's which explain his views about the wrongness of slavery (c). If the features upon which the putative wrongness of slavery supervene were very different, then Quincy wouldn't think that slavery is wrong. So even though the best explanation of Quincy's view about the wrongness of slavery does not *refer* to any moral properties, the fact that

[18] Further, were Bobbi to accept that this was the cause of her view, she likely would no longer feel so confident about homosexuality's wrongness. This, however, is not a necessary part of the Hurley-Sturgeon argument. See Nozick 1981, 348-352 for more on this. See Geuss 1981 for more on the structure of this problem as it is faced by Critical Theory, especially p. 61. I will briefly discuss this later.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Quincy's view is explained by the features upon which the wrongness of slavery putatively supervenes preserves the possibility that Quincy knows that slavery is wrong. We might reasonably think that the feedback condition (the presumptuous thought) is thereby met, and so that there are moral facts.[19]

Some explanations that do not refer to moral properties (the homosexuality example) undermine the moral views they explain, while others (the slavery example) corroborate the moral views they explain. What seems important is not whether the explanations of our moral attitudes refer to moral properties, but whether they are sensitive to the features upon which morality supervenes. Thus the supervenience of the moral upon the nonmoral seems to open up the possibility that one can know moral facts by knowing which nonmoral features of the situation the moral features supervene upon, knowing whether those nonmoral features obtain, and having one's moral views sensitive to this latter bit of knowledge. Harman's basic argument, then, is apparently invalid.

But one might think that the supervenience thesis deceptively gives the moral objectivist what he hasn't actually earned. To see how so, let us consider a ridiculous example in which the supervening properties obviously do not merit an objectivist understanding.

[19] Further, Quincy can accept the explanation that his belief about the wrongness of slavery is explained by his beliefs about the effects of slavery, without the former belief being thereby debunked. See previous footnote.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Suppose Knute is attracted to any woman over 6 feet tall. The best explanation for why Knute is attracted to such people very likely does not refer to some property of attractiveness possessed by them, even though that is how Knute is wont to describe them.

But *ex hypothesi* Knute's beliefs about the attractiveness of such people are supervenient upon a straightforwardly physical property they possess, namely, their height. So let us suppose that Knute thinks that Heidi, who is over 6 feet tall, is attractive. Further we can suppose that if Heidi were not over 6 feet tall—she might rapidly experience the effects of aging—Knute would no longer think that Heidi is attractive. Further we can suppose that Knute is fully aware that his belief about Heidi's attractiveness depends upon his belief about her height. So Knute knows that Heidi's attractiveness supervenes on her height, knows whether Heidi is over 6 feet tall, and wouldn't think that Heidi was attractive were she not over 6 feet tall. Knute's belief in Heidi's attractiveness seems to depend upon facts about Heidi's attractiveness no less than Selena's belief about the wrongness of burning the cat depends upon facts about the wrongness of burning cats. If Selena has moral knowledge, Knute has knowledge about who is attractive.

But, of course, this is absurd. The mere fact that Knute's beliefs about Heidi's attractiveness supervene upon physical properties is no insulation from the kind of anti-objectivist interpretation we would naturally make. Somewhere we have gone awry.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Harman identifies the problem we face as follows:

> We can distinguish between two related propositions: (1) Features of acts that make the acts wrong sometimes explain things that can be observed. (2) The fact that certain features make acts wrong sometimes explains things that can be observed.[20]

Features of acts that make them wrong—or features of persons that make them attractive to Knute—can indeed explain things that can be observed, like Selena's belief that those acts are wrong, or Knute's beliefs about who is attractive. But that is not what is needed to establish an objectivist interpretation of the domain in question. Objectivism requires that *the fact that* certain features make acts wrong—or the fact that certain features make persons attractive—explains things like Selena's observation that *m*, or Knute's belief that Heidi is attractive. The physicist we met before thinks she understands how the fact that there is a proton in the cloud chamber could explain her seeing what she observed. The relevant parallel question is whether we think that the wrongness of some action explains Selena's thinking it to be wrong, and the possibility that this explanation could take place via the supervened-upon features does not relieve us from the daunting task of understanding how the *wrongness* of an action could explain anything, whether it be Selena's belief or the presence of the supervened-upon features. Thus it seems we need to understand the

[20] Harman 1986, 63.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mechanism whereby our moral views are influenced by the rightness and wrongness, goodness and badness of things in the world:

> What's needed is some account of *how* the actual wrongness of [someone's] action could help to explain [someone's] disapproval of it.[21]

Without such an understanding, the claim that one knows about some matter, even if it supervenes on some more securely objective realm, fails to guarantee the correctness of an objectivist understanding of it.

But we need to be on guard against construing the feedback requirement in a way that grants the anti-objectivist too much, that places the objectivist in an unnecessarily tight corner. At a minimum, Hurley has demonstrated that our moral beliefs can track 'moral truths'. Harman argues that counterfactual tracking is not enough to establish moral objectivity—we need to understand the mechanism connecting moral features to the things we uncontroversially observe. By focusing on understanding *how* we know about right and wrong, it might appear that Harman is insisting that moral truths must cause our moral observations. And while our original working assumptions seemed reasonable to grant, it is not fair to saddle the moral objectivists with the task of showing that moral truths are *causally* efficacious.[22] For

---

[21] Harman 1986, 63.

[22] I should note that plenty of moral objectivists do think that moral properties cause (many of) our moral beliefs, where the operative notion of cause is efficient causation. If true, this would support the objectivity of morality. But many objectivists do not hold this view, and I, for one, rely on no such thought. Philosophers who think that only causal explanatory efficacy can underwrite realism—such as Nancy Cartwright—also do not think that theoretical physics is to be understood realistically. See Cartwright 1983. The interesting question, however, is whether there is a contrast between the objectivity of mathematics and/or theoretical physics on the one hand, and morality on the other.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

instance, no one doubts that it would be a minor victory for the moral objectivist if he could show just that morality and arithmetic merited similar interpretations. And yet it is obvious, as even Harman himself admits, that we do not even understand what it would be like to be in causal contact with arithmetical objects.[23] So, in parallel fashion, the success of moral objectivism ought not ride on whether moral properties or truths cause anything.

The lesson to be learned from the failure of Hurley's account is that we cannot substitute for a clause occurring in an explanation another clause that always happens to have the same truth-value. If I believe that the kids are doing something wrong in part because I believe that they are igniting the cat, and if the kids are doing something wrong if and only if they are igniting the cat, it does not follow that I believe that the kids are doing something wrong because they are (indeed) doing something wrong. The "because . . ." context is not necessarily extensional; it does not always permit this kind of substitution.[24]

To summarize, we have found that supervenience alone does not provide morality with the objectivist foundation we were seeking, but the expectation that the objectivity of morality depends on its causal efficacy threatens to subjectivize more

---

[23] Harman 1977, 10.

[24] Some 'because . . .' contexts *may be* extensional, e.g. those of efficient causation. But, as I just said, we have been given no reason to think that morality must enter into *efficient-causal* explanations in order to construe it realistically, and, also, there are good arguments to the effect that even these kinds of explanations do not provide extensional contexts See "Causality and Extensionality" in Anscombe 1981, especially 173-179.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

than just morality. We need a test for the objectivity of morality stronger than the Hurley/Sturgeon supervenience test, and yet weaker than the causal efficacy test.

Reflecting further on the nature of arithmetic may suggest a middle ground. Arithmetic survives Harman's test, because while mathematics does not causally explain anything we observe, mathematical principles often figure in explanations of what we observe, and so there is a kind of observational evidence for mathematics. That is, the best explanations of the observations that scientists concern themselves with make use of mathematical statements. Mathematical statements can indispensably figure in explanations of observations, even though the referents of the mathematical concepts contained in such statements possess no causal properties. In other work, Harman has also recognized the legitimacy of rational explanations, of statistical explanations, and of relying upon generalizations to explain particulars.[25] Causal explanation, it seems, is not the only relevant kind of explanation that can confer objectivity.

Perhaps morality could occupy a similar position. The moral objectivist may grant that moral truths do not causally explain anything, and still maintain that the best explanation of Quincy's view that slavery is wrong involves the fact that slavery *is* wrong. And the anti-objectivist can bicker with the truth or the power of this explanation, even if she grants that Quincy can track the properties upon which the

---

[25] Harman 1973.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

wrongness of slavery supervenes. The issue dividing them is not whether moral truths causally explain our moral observations, but whether the best explanation of our moral observations, causal or no, indispensably refers to the content of those observations. Thus the reference-in-explanation test, as I originally formulated it, seems to be the best way to understand what divides the objectivist from the anti-objectivist.

Once the anti-objectivist grants that the type of explanation in question need not be causal, even the most unscientistic moral objectivist should be willing to grant that the reference-in-explanation test is appropriate. Consider the following two passages written by Thomas Nagel, an unscientistic objectivist if there ever was one. The first passage is from "The Limits of Objectivity", Nagel's 1980 Tanner Lectures. The latter passage is from *The View from Nowhere*, his 1986 book a large part of which is a revision of his 1980 lectures.

> It begs the question to assume that explanatory necessity is the test of reality in this area. The claim that certain reasons exist is a normative claim, not a claim about the best explanation of anything. To assume that only what has to be included in the best explanatory picture of the world is real, is to assume that there are no irreducibly normative truths. [26]

> Mackie meant that reasons play no role in causal explanations. But it begs the question to assume that *this sort of* explanatory necessity is the test of reality for values. The claim that certain reasons exist is a normative claim, not a claim about the best *causal* explanation of anything. To assume that

[26] Nagel 1980, 114.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> only what has to be included in the best *causal theory* of the world is real, is to assume that there are no irreducibly normative truths.[27] [*Italics* mine.]

Apparently, during the six years between the writing of these two passages, Nagel came to see that while the objectivity of morality need not be bound up with the causal efficacy of moral properties—which most defenders of the objectivity of mathematics would grant—there are nonetheless different modes of explanation that morality fits into if it is objective. So long as we are flexible in our understanding of what an explanation can be, the reference-in-explanation test does not look unreasonable. Harman's original argument appears valid after all.

2.5 Objectivity and Justification

Let us now consider a different sort of attack on the reference-in-explanation test toward which the moral objectivist might now be tempted. Rather than directly denying that morality needs to fulfill a certain explanatory role in order to be understood objectively, the approach I now want to consider argues "that the existence of moral explanations cannot ground or justify morality, so that a fundamental skeptical position about morality remains open even if we grant their existence."[28] On this view, whether there are successful moral explanations turns out

[27] Nagel 1986, 144.

[28] Copp 1990, 239.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to be incidental to whether we are justified in acting one way rather than another. Thus, it is argued, the reference-in-explanation test is insufficient for establishing the objectivity of morality.

To see the plausibility of this position, consider explanations that refer to concepts belonging to moral perspectives we reject:

> A Nietzschean might seek to explain Stalin's ruthless behavior on the basis that he was an (approximation to the) overman. But this would be an ordinary psychological explanation that incidentally invoked a Nietzschean moral concept. Psychology could even adopt the Nietzschean concept and postulate the existence of an "overman" personality, but we would not be tempted in the least to accept the explanatory utility of the concept as justifying any standard that treats being overmanlike as a virtue or ideal.[29]

Even though some moral views consider the overman to be the ideal character type, the fact that some explanations of our observations might refer to overmen does not go any way toward justifying such views. The explanation's reference to the overman is incidental to the justification of any moral standard; accepting such explanations is compatible with rejecting such standards.

Much the same can be said for concepts such as "good" and "right". We might find it useful to explain someone's action by saying that he is a good person. It may be that the term "a good person" picks out some complex psychological disposition, and this disposition is one that can illuminate otherwise inexplicable patterns of action. Grammar school children often use the term in this way, referring

[29] Copp 1990. 247-248.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to "the good kids" and "the bad kids". But one who accepts such explanations may still coherently question whether it is unjustifiable to fail to be a good person.

There is not much about this argument I want to quarrel with. It is one thing to accept that a property is explanatorily useful; it is quite another to accept that such a property constitutes a justificatory standard. But we shouldn't infer from the fact that the explanatory usefulness of a property is not sufficient for a certain kind of justification that the explanatory usefulness of a property is not necessary for that kind of justification. It may be the case that we are free to reject as a normative standard a moral concept we use in explanations, but this does not license the claim that the normative standards we do adopt play no explanatory role. Normative standards may not derive any of their justificatory force from their explanatory capacity, but that would not tell against the link between explanation and moral objectivity I have been defending.

2.6 The Objectivity of Artifacts

As we have seen, if we use the reference-in-explanation test as a criterion for objectivity, much of what we ordinarily think of as objective—science and mathematics—likely passes the test, while much of what we usually take to be subjective—sexual attractiveness—likely fails it. Thus it seems that the reference-in-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

explanation test is a good way to determine whether an objective understanding of more controversial realms is appropriate.

Warren Quinn, however, has pointed out that the reference-in-explanation test rules out a couple of areas we would otherwise think merit an objective interpretation.[30] The test not only casts doubt on the objectivity of morality, but also on the objectivity of many of the classifications of ordinary life not used in the sciences, including natural objects (e.g. gardens) and artifacts (e.g. chairs). What observation is explained by the putative fact that my garden is full of weeds? What observation does the putative fact that the object I am sitting on is a chair explain? The physical parameters of things the chair affects are best explained not by the fact that it is a *chair*, but by its shape, size, mass, density, opacity, and so on. We could just as well forget to mention that it is a chair. Similarly, the ways in which the chair supposedly affects my actions, like my choosing to sit on it, my giving the cashier money for it, could also be explained by showing how these physical parameters interact with my "theory" of chairs. All necessary mentioning of chairs takes place within intensional contexts; an explainer of my thoughts and actions needn't use the concept of chair herself in order to discharge her task. The fact that the thing I am sitting on is a chair, if it is a fact, seems to explain nothing.

[30] See his "Truth and Explanation in Ethics" in Quinn 1993.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Quinn thinks that this argument functions as a *reductio*, since we are more certain about the existence of chairs than we are about the legitimacy of the reference-in-explanation test. But there is a gap between holding a philosophical view about the existence of artifacts, and the question of whether any of our thoughts concerning artifacts are appropriate. Not every thought that fails the test for objectivity should be expunged; simply consider your thoughts about the attractiveness of your spouse (or your would-be paramour). It is doubtful that you are able to detect this property of the world—his/her attractiveness—with respect to which many other people are cognitively impaired. You are not wrong to think that your spouse is attractive; but you wouldn't necessarily be wrong if you thought otherwise either. The correctness of your judgment—insofar as we can speak here of correctness—depends not only on the qualities possessed by your spouse, but also, I suppose, by your own feelings, thoughts, and the like.

Similarly, the anti-objectivist does not necessarily think we should give up talking about sexually attractive people, chairs, gardens, or morality. Rather, the appropriateness of using these concepts stems from our desires, purposes, projects and the like, none of which aid us noninstrumentally in knowing the features of the world. (Noninstrumentally, because we may need certain desires—like the desire to spend significant time in the laboratory—in order to gain some pieces of knowledge. But the desire itself does not put us into cognitive contact with the world.) While

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Quinn is right to think any philosophical view foolish that instructs us to give up our views about ordinary objects like chairs and gardens, it is not wrong to see the concepts we use to denote these objects as dependent on and responsive to the projects we pursue. Someone who failed to recognize the objects we think of as chairs as a distinct kind, having only a concept, say, for an object to be sat on (thus failing to distinguish chairs from sofas), wouldn't be suffering from any kind of cognitive failure. He would simply not live as we do.

2.7 THE ABSOLUTE CONCEPTION

We have seen that the reference-in-explanation test survives attacks based on the supervenience of the moral, on the gap between explanation and justification, and on the nature of artifacts. Finally, I want to defend the validity of the test from attacks based on *the intelligibility of the explanandum*, attacks put forward most forcefully by John McDowell. Earlier I formulated the reference-in-explanation test in the following way: if the best explanation of (some of) our moral observations need make no reference to moral facts, moral objectivism is false. But perhaps a more appropriate test for the objectivity of morality would ask whether she who fully explained our moral observations could nevertheless always deny what they say.[31] Even if the best explanation of our moral observations does not refer to moral facts,

[31] McDowell 1985.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

perhaps one cannot grasp the content of those observations well enough to explain them if one also denies their claims about how things are.

Let us take a fresh look at the relation between the reference-in-explanation test and objectivity. Start with the very simple idea that there can be a difference between the way a thing appears to be, and the way it really is. John McDowell illustrates this distinction by pointing out the way we correct for the angle from which we observe a plane surface when we judge its true shape.[32] A plane surface may appear to be trapezoidal to one who is situated close to and just above the surface, looking along the surface from one edge. But we (unthinkingly) correct for the point of view from which we view the surface, judging it to be in fact, say, rectangular. The thought that the surface is rectangular reflects the way the world really is, as opposed to the thought that it is trapezoidal, being only a mere appearance.

Part of what validates the claim that the surface is in fact rectangular is that it can be deployed in conjunction with other facts (e.g. those about optics, geometry, physiology, and the like) to explain why the surface appears trapezoidal from one particular point of view. If all we had was the surface appearing trapezoidal from one point of view, and appearing rectangular from another, with no way to relate the two, we would have no firm ground for taking either of these appearances to reflect the way the world really is. But since we can explain why the surface appears trapezoidal

---

[32] McDowell 1983.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

on the basis of what else we know about the world, including the observation that it appears rectangular from this other point of view, we are licensed to judge that the surface is in fact rectangular, and that its appearance as trapezoidal from the first point of view is only an appearance and no more. The best explanation makes no reference to the surface *being* trapezoidal.

This ordinary use of the distinction between the way a thing (merely) appears and the way it really is readily suggests numerous metaphorical extensions the debunker of objectivisms can employ. For instance, we can extend the notion of a point of view to the visual apparatus humans possess. So from a particular point of view, one constituted by the possession of the human visual apparatus, things appear a certain way—colored. Then we can marshal a scientific argument to show that things appear colored to perceivers like us because of the microstructure of the object's surface, the nature of the light, and the details of the visual apparatus with which we are equipped. The appearance that things are colored gets explained by these other facts. Thus it seems that one could deny of any particular object (e.g. this piece of paper) that it is colored (white) while still explaining why it would appear colored (white) to observers like you and me. That objects are colored is not an objective fact, then, for one could explain why they appear that way without oneself being taken in by the appearance. The world conceived objectively contains no colors at all.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Now what has this to do with the objectivity of ethical value? Well, if we could make sense of extending the notion of a point of view to the human visual apparatus, then it seems natural we could extend it to something like the interests and moral sensibilities that characterize a particular ethical outlook. From a particular ethical outlook, then, certain courses of action (say) will appear to have an evaluative character, a character not perceptible from the standpoint of other rival ethical outlooks. The opponent of moral objectivism, then, hopes to explain these appearances of value from some more fundamental standpoint, one which illuminates why certain things appear important to denizens of a particular ethical outlook, but not to others, all the while committing herself to none. If successful, she would have shown that ethical values do not belong to the world as it is in itself, that ethical value is not objective.

Bernard Williams argues for the possibility of just such an explanation of our valuations.[33] He introduces the notion of the 'absolute conception of the world', a notion that exploits and extends our ordinary understanding of the contrast between how a thing appears to be and how it really is, the understanding we met just a moment ago when considering how the shapes of things appear. The absolute conception of the world is the single conception of both the world and the multiple points of view occupiable within it. If all intelligent enquirers converge on a

[33] Williams 1978, 241-246. See also Williams 1985, 132-155.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conception of what the world is like, and the best explanation of their one mind involves the idea that things are as this conception represents them, then and only then may we say that their conception is objective.

Two things are to be noted about the absolute conception. First, in order to earn the lofty title of being the absolute conception *of the world*, this conception must not simply be what all of these *Überwissenschaftleren* jointly think, but must also explain why others have different thoughts about the world. The absolute conception must account for why things appear the way they do both to the *Überwissenschaftleren* and to us mere *Menschen*. Otherwise, it is nothing more than one conception among others. (Recall the analogous problem we encountered above if we have no way to relate the point of view from which the surface appears rectangular with the point of view from which it appears trapezoidal.)

Second, it is possible that the best explanation of their convergence on a thought might *not* rely on this answerability to the way things are. If all of these *Überwissenschaftleren* preferred split-level suburban homes to urban lofts, the best explanation of this convergence would most likely not involve the idea that split-level homes really were the better of the two. A more powerful explanation of their preference probably lies in the psychological and social sciences—in a domain other than architectural aesthetics—and, presumably, the convergers would happily be on to *that* explanation of their own preference as well. So not every thought converged

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

upon is objective; only those which figure into the explanation of the convergence, as well as those which explain why others have different thoughts, earn that status.

Williams characterizes the work done by the absolute conception this way:

> The substance of the absolute conception . . . lies in the idea that it could nonvacuously explain how it itself, and the various perspectival views of the world, are possible.[34]

Note how Williams' use of the absolute conception of the world to determine whether something is objective is a natural articulation and extension of the reference-in-explanation test. It is an articulation of the reference-in-explanation test, for the absolute conception must be able to explain how it itself is possible using the concepts and content it already contains. It is an extension of the reference-in-explanation test, for the absolute conception provides some formal criteria for determining whether one explanation is indeed better than another. The explanation that seems most natural from one standpoint—that the plane surface appears trapezoidal because it *is* trapezoidal—is displaced by the explanation that issues from the conjunction of all standpoints—that the plane surface appears trapezoidal because it is rectangular and rectangular objects viewed from standpoints close to the surface and at about the same height as the surface, looking along the surface from one edge, appear trapezoidal. We can see that the latter explanation displaces the former because it satisfactorily explains why we might have thought that the surface was

[34] Williams 1985, 139.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

trapezoidal from the one point of view. But the plane surface is not trapezoidal: the best explanation for its appearing to be trapezoidal does not involve the thought that it is trapezoidal. So the absolute conception endorses the basic insight of the reference-in-explanation test—that a truth is objective if it figures in the best explanation of how things (including moral observations) are—while explicating how to tell whether one explanation is better than another.

Now Williams thinks that our ethical thoughts do not have a place in the absolute conception of the world, and thus that there are not objective ethical truths. It is not that he doubts that convergence is possible in ethics, but rather that the best explanation for any such convergence would rely on the social sciences, i.e. on a domain of thought *other* than ethics.[35] Like the thought about split-level homes being more suitable than urban lofts, our moral thoughts can be adequately understood and explained in ways that do not affirm the value of what we find important. The social scientific theory that successfully accounts for our moral observations, then, would be conceptually rich enough to explain why people evaluate the way they do, while evaluatively sterile enough to avoid reinstating the objectivity of a particular set of ethical values. The existence of such a theory would undermine ethical value's claim to be objective, showing ethical value to be no more than how things appear for

[35] Williams 1985, 150.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

people living in *one* kind of social world, which presumably is not the only kind of social world someone can inhabit.

McDowell, however, has reservations about whether the absolute conception of the world could have this kind of explanatory power. This is not because he doubts the intellectual prowess of present-day and future scientists. Rather, he thinks there are conceptual problems with trying to get a hold on the phenomenology of experience from a standpoint that has completely washed its hands of the details that characterize that experience. He rhetorically queries:

> Can the expansion to embrace the various local points of view be undertaken in the objective spirit that would be required for its upshot to sustain the correlation between objectivity and reality? Or would it necessitate—surely defeating the project—a regression from the attempt to transcend particular points of view, in order to achieve an undistorted picture of reality as it is in itself, to an unregenerate occupation of the points of view that were to be transcended?[36]

Before assessing whether the absolute conception of the world might explain our valuations, McDowell first illustrates the problems awaiting the debunker of the objectivity of secondary qualities such as color, qualities that are clearly not part of the absolute conception. Williams himself admits that the absolute conception somehow needs to explain " '. . . looks green' in some way that does not presuppose any prior understanding of '. . . is green.' "[37] It is hard to see how one could even grasp the meaning of the explanandum—that some things appear green—without

[36] McDowell 1983, 10.

[37] Williams 1978, 243.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

implicitly relying on a standpoint from which colors present themselves as properties things have. To transcend all points of view from which one sees colored objects is to render *what is to be explained* unintelligible. McDowell writes

> [I]t seems to be an illusion to suppose that such understanding could still be forthcoming after we had definitively left behind a view of the world which represents colours as properties that things have . . .[38]

If we are to debunk the objectivity of color, we will need both to occupy a point of view from which we can understand the color predicates as they occur in the description of the content of experiences of (putatively) colored objects, and yet to stand clear of any point of view from which they are used to state how things are. But this is a position the possibility of which is difficult even to conceive, much less occupy.

An example may prove helpful here.[39] It just so happens that if you take a green object and shine red light on it, it will appear black. Now this is fully intelligible only because we already know what it is to say of an object that it *is* black. Without that prior understanding, we would be at a loss to say just *how* the object appears under red light. To be sure, from the standpoint of the absolute conception one could explain someone's inclination to utter the word "black" when confronted with something with a certain microstructure. But this is not to explain why things

[38] McDowell 1978, 10.

[39] I owe this example to Dan Garber, who used a variant of it to critique Berkeley for failing to distinguish appearance from reality. Any mistaken inferences drawn from the example are wholly my responsibility.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

appear black to those with standard human visual equipment. Describing the content of the appearance is the prerequisite for the explanatory task of the absolute conception of the world, and this is why one cannot grasp how things might appeared colored without also thinking that some things in fact are colored. So even though the color of an object is a subjective property, one is in no position to explain away the reality of color by invoking this fact. Subjective properties cannot be pried apart from their objects as easily as one might have thought. Let us call this conceptual connection between appearance and objectivity that McDowell points to *the connectedness thesis*.

The connectedness thesis has important implications for the debunker of moral objectivity excited by the standpoint framed by Williams' absolute conception of the world. McDowell readily grants that moral values are not something objective in the sense identified by Williams. Like colors, moral values are conceivable only from a standpoint wrapped up with the idea of an experience of an object's seeming to have it: moral values are hardly anything like primary qualities. And yet this fails to impugn their objectivity, for they cannot be explained from a point of view not already committed to them. In order even to grasp that it is appearances of *moral value* that we are to be explaining, we thereby commit ourselves to the claim that at least *something* is morally valuable. The absolute conception of the world, then, is a fiction, for there is no point of view from which both to explain appearances of moral value and to steer clear of thinking of moral values as objective.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

McDowell's connectedness thesis is compelling. It is difficult to conceive of a situation where one would say that something only appears F, without being willing to say of any other object whatsoever that it *is* F, where F is some phenomenal predicate. That we have the predicate in our vocabulary suggests that it is appropriately predicated of some object somewhere. Perhaps we could come up with peculiar examples of phenomenal predicates for which this does not hold true, but it is more doubtful we could for an entire range of predicates, such as those used in "evaluative discourse".

There are a host of concerns one could raise about the connectedness thesis, most of which I cannot discuss now. There is one concern I do want to mention, however, in order to set it aside. The cultural anthropologist may dispute McDowell's claim that one could not understand how things can appear to another to have certain properties, after having left behind a view of the world which represents objects as having those properties. So might the historian. The thesis that one must buy into a world-view in order to understand how its predicates get deployed seems to eliminate the possibility of understanding the world-views of those very different from one's own. It even seems to eliminate the possibility that one can *consider* adopting a world-view much different from one's own. A leap of faith, then, seems to be the only way to get oneself into understanding a different outlook, and, by then, one has already adopted it. The role of reason in weighing the competing merits of various outlooks has closed up. Perhaps reason can direct us in applying whatever concepts

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

we find ourselves to be using. But it cannot help us decide which concepts are appropriate ones with which to understand the world.

This argument against McDowell is much too quick. I raise it only to show some problems that could be considered. I will not directly pursue this line of argument, for I think these worries can be more fruitfully addressed via another route.

Let us consider granting McDowell the connectedness thesis. So perhaps I am in no position to hold that something appears valuable to someone else without thinking for myself that something is valuable. This instantiation of the connectedness thesis certainly does have a ring of plausibility to it.

It remains an open question, however, whether the connectedness thesis is *relevant* to the status of the objectivity of moral valuations. The importance of the things one finds valuable is not that they produce certain sorts of appearances, images, qualia, and feelings in a person, but that they guide one's action in particular ways. Values, whatever they are, are not phenomenal in the way colors are. Valuing something and seeing an appearance of something are two different kinds of things. This means that valuation need not (and ought not) be understood as an *appearance of* anything, since different persons can take the same appearances as good grounds for very different courses of action. Perhaps there is some link between how things appear to one and what one values, a link tight enough for the connectedness thesis to become relevant. But we would need some sort of argument for such a thesis, and no argument has been given nor is obvious.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ethical thought construed this way escapes whatever conclusions are forced upon us by McDowell's connectedness thesis. There may be a conceptual connection between appearance and reality, but since it is far from obvious that valuations are to be thought of as appearances, or that they are tightly grounded upon appearances, this connection does not seem relevant to understanding whether explanations of valuations can be grasped by those who do not share them. McDowell himself later recognized this feature of ethical thought, admitting that values are not in fact essentially phenomenal qualities.[40] Indeed, they are not phenomenal at all. But without the support of the conceptual connection between appearance and reality that McDowell forges, it is hard to see what makes it so difficult for one to understand others' moral evaluations, understand them even though one doesn't share them. And so we have no good reason to think that we cannot explain a person's moral observations without in some way admitting that things are as she thinks. The reference-in-explanation test—and its extension as the test to see whether moral truths are part of the absolute conception of the world—still seems to be the best way to articulate the feedback requirement (the presumptuous thought).

---

[40] McDowell 1985, 175.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2.8 THE NEED FOR A THEORY OF ERROR

A more promising line McDowell considers is whether the anti-objectivist can understand not the simpler psychological states that constitute moral valuation, but the more complex psychological states that guide the virtuous person to make decisions when multiple morally relevant considerations are operative. In "Virtue and Reason", McDowell adduces considerations intended to make plausible the Socratic identification of virtue with knowledge.[41] The virtuous person realizes that certain sorts of situations require certain sorts of responses from her, and the Socratic thesis identifies this realization with a kind of knowledge had by the virtuous person. In order for an action to count as a product of virtue, this realization must fully constitute the agent's reason for acting as she does. If she also requires some extra incentive in order to comply with the requirement, she fails to act virtuously. Virtuous action must be motivated by considerations which operate categorically.

What this position excludes is the possibility that someone has the same cognitive states had by the virtuous person and yet fails to act as the virtuous person does. This possibility would open the door for someone to describe the cognitive psychological states responsible for the virtuous person's action, while resisting the conclusion that the action is required, this because the explainer lacked the extra incentive had by the virtuous person. In such a case, one could explain the virtuous

[41] McDowell 1979.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

person's motivation—that is, cite the desire she had—without endorsing her action. This possibility is one which the anti-objectivist excited by the absolute conception relies upon in order to discharge his task. Only if he gets a purchase on the psychological states of the virtuous person can he explain her action, one of the necessary conditions for any conception of the world to qualify as absolute.

McDowell is prepared to deny that the anti-objectivist can do this. The virtuous person's cognitive makeup is supposedly one that a nonvirtuous person cannot fully share. A nonvirtuous person's view of the situation is clouded by "obstructive factors, for instance distracting desires."[42] These obstructive factors prevent the nonvirtuous person from understanding her situation accurately, which for McDowell is both the necessary and the sufficient condition for consistently understanding and performing virtuous action.

Traditionally, ethical noncognitivists have maintained that all action, including morally-motivated action, is produced not only by a belief about the world, but also by some desire. This way, someone could share the same belief about the world had by the virtuous person, while lacking the motivation to act morally, for she might lack the appropriate desire. McDowell alters the picture in an interesting way. He agrees that the virtuous person and the nonvirtuous person have different desires, and that this difference is in some sense responsible for the difference in their actions.

---

[42] McDowell 1979, 348.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But McDowell preserves the cognitivity of virtue by suggesting that the nonvirtuous person's desires prevent her from sharing the beliefs had by the virtuous person. Desires influence action through the medium of beliefs, and so McDowell rules out the possibility that one could share the virtuous person's conception of a situation while having different desires.

Now earlier we saw that McDowell argued that one could not understand how the value of things appears without recognizing the reality of value oneself. We also saw that the language of appearance is one that the anti-objectivist needn't buy into, and so that McDowell's argument misses it mark. In "Virtue and Reason", McDowell also holds that "it is highly implausible that all the concerns which motivate virtuous actions are intelligible, one by one, independently of appreciating a virtuous person's distinctive way of seeing situations."[43]

Fortunately, this is not the main thrust of his paper. Here McDowell thinks the anti-objectivist's real trouble arises when he tries to account for why the virtuous person selects to act on one concern *rather than* another.[44] The thing which explains this selection is her "conception of how to live", a conception that guides a person as she confronts various incompatible demands on her. Her conception of how to live identifies what she finds to be most important. Her conception of how to live supplements the "core explanation" of her action, supplements the thing that explains

---

[43] McDowell 1979, 346.

[44] McDowell 1979, 344-346.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that such-and-such is a relevant concern. Even if the anti-objectivist can grasp the core explanation for why the virtuous person does what she does, he cannot, from the outside, grasp her conception of how to live, much less explain it.[45] The anti-objectivist cannot understand why some concern of hers is the salient one unless he himself sees it "as a reason for acting which silences all others", as a reason in the presence of which other considerations have none of their normal reason-giving force. Thus he cannot show that the best explanation of her moral observations fails to make reference to moral facts.

McDowell denies that one can grasp the virtuous person's perception of which concern is salient without also seeing for oneself that it is salient. Why? He thinks this intelligibility constraint on one's conception of how to live follows from the *uncodifiability* of one's conception of how to live.[46] One cannot explicitly formulate in universal principles one's conception of how to live,[47] nor can one lexically rank one's concerns such that conflicts between them could be resolved merely by consulting such a list.[48] Following Aristotle, McDowell thinks that generalizations about practical matters hold only for the most part, and that there frequently are exceptions to the best generalizations, exceptions which cannot be antecedently

[45] McDowell 1979, 345.

[46] McDowell 1979, 345.

[47] McDowell 1979, 343.

[48] McDowell 1979, 344.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

accounted for. Only someone who held for himself such a conception of how to live could anticipate such exceptions.

I have no qualms about thinking of a conception of how to live as essentially uncodifiable. Indeed, it seems to make a good deal of sense. But we may wonder about the link McDowell wants to forge between the uncodifiability of a conception of how to live and the external unintelligibility of a conception of how to live. Why is uncodifiability sufficient for external unintelligibility?

This is something about which McDowell is not explicit. It appears to do with his discussion of Wittgenstein earlier in his paper. Specifically, McDowell argues that it is an illusion to think that we reason in ways not dependent "on our partially shared 'whirl of organism'."[49] The fact that human nature is common and that our forms of life are shared makes it possible for us to understand one another. There is no understanding or displaying our rationality from a position completely outside our form of life.

As before with Aristotle, I have no problem with McDowell's understanding of Wittgenstein: it makes sense to think that sharing the same *Lebensform* is necessary for intelligibility and communication. What does cause this reader concern is the propriety of McDowell's application of these Wittgensteinian insights to the question of the external intelligibility of conceptions about how to live. Even though

---

[49] McDowell 1979, 340.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

someone else's conception of how to live is uncodifiable, and even though I do not fully share her conception of how to live, I probably can with success predict which of her concerns she will find "salient" in various situations, reliably guess what kinds of things she deems most important. Intuitively, it seems flatly false that persons must share the same conception about how to live in order to understand each other's conception. Part of what is interesting about talking to (and living with) others is discovering the *differences* among us. We couldn't discover such differences if we couldn't understand one another.

But perhaps McDowell is thinking of more stringent standards of intelligibility. It could be that a conception of how to live counts as intelligible only if one could predict perfectly which concerns someone holding that conception will find salient in various situations. This, to be sure, I could not do for someone who held a conception of how to live different from mine. But it is doubtful I could do this for anyone (even myself?), whether we shared the same conception or no. Of course, if the criterion for judging the identity of conceptions just is the standard of perfect prediction, McDowell's claim is trivially true. But if it is not, the claim seems empirically false.

Let us ignore these problems, and assume McDowell is somehow right about uncodifiability being sufficient for external unintelligibility. This would support his thesis that one cannot grasp the virtuous person's conception of how to live unless one

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is oneself virtuous. One would be unable to explain the virtuous person's valuations without valuing those things himself.

Unfortunately (for McDowell), that is not the only consequence of this position. The Aristotelian considerations that McDowell cites would then support not only the view that the virtuous person's conception of how to live is uncodifiable, but also the view that the *vicious* person's (or indeed any person's) conception of how to live is, in most cases, likewise uncodifiable. And if applications of Wittgensteinian insights to uncodifiable conceptions of how to live support the external unintelligibility of virtuous action, then they support the external unintelligibility of *vicious* action as well. This has the unhappy consequence that one cannot explain someone's vicious action unless one is oneself vicious, and vicious in the same way. *No one's* actions are fully understandable, unless we share their conception of how to live. This, I think, is an odd result for an argument informed by Wittgenstein. Such unintuitive conclusions demand that we reject the premise that uncodifiability implies external unintelligibility.

Let us try to rescue McDowell one more time. Perhaps, for some reason not yet in view, the virtuous person's conception of how to live is fully intelligible only to other virtuous persons, though it remains possible that the nonvirtuous person's conception of how to live is intelligible to the virtuous. Plato in fact appeared to hold

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

something like this view.[50] Now, what reason could McDowell have for thinking that such a thing is true?

Here is one option, the only one I can think of: John McDowell himself—or anyone else backing this position—is a virtuous person, and *qua* virtuous person he can fully understand the conceptions of how to live of both virtuous and nonvirtuous people, and he knows he can do this. If we have our doubts about whether the virtuous person is in this unique position, it is because we are not virtuous as well.

The sharp distinction between ethics and metaethics has been repeatedly attacked by moral philosophers in the last thirty years, but the view now being considered would practically *identify* the two. One must be virtuous in order to answer questions about moral objectivism, about the nature of moral disagreement, and so on. I don't mean to scoff at this possibility—it very well could be true. But it is hard to swallow without some argument that we mere *Menschen* (and not only *die Tugendbolde*) can understand. To do so would require thinking of moral philosophers as a kind of priest, one who is able to tell the rest of us both, say, what the conceptual relation between nature and morality is, *and* what is actually of value, though not able to offer an argument for either that we can grasp.

This is a point Bernard Williams has repeatedly emphasized. When discussing the views of McDowell, Williams frequently complains that the claims of

[50] Plato, *Republic*, 409de.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the moral objectivist are question-begging until we also have *a general practicable theory of error*.[51] We need to "account generally for the tendency of people to have what, according to its principles, are wrong beliefs."[52] We need to understand why our conceptions of how to live are frequently out of step with the virtuous person, if we are to accept the thought that our conceptions are sometimes wrong, rather than merely different. McDowell has taken the first step toward this end, by suggesting that failure to perceive what is of value often stems from a distracting desire which blurs our appreciation. Wishful thinking often prevents one from seeing things aright. But this suggestion is only a beginning, and by itself, inadequate. Failure to offer an adequate theory of error, or something else that fills the role of such a theory, results in the deformation of morality Williams calls *moralism*:

> The insistence that a given person is wrong, disconnected from any possible understanding of how it comes about that he is wrong, tends to leave the commentator entirely outside that person, preaching at him.[53]

An argument such as McDowell's needs to convince us that *there is* a single common truth about which we might not yet be on to, if it is to convince us that immoral views are not only immoral, but mistaken. But this the moral objectivist has not done. And as long as he has not done it, the insistence that those whose are not virtuous are somehow making a mistake is, as Williams puts it, "mere bluff". (In

---

[51] Williams 1995a, 182-191.

[52] Williams 1985, 151.

[53] Williams 1985, 219n16.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

more recent writings, McDowell admits that something akin to religious conversion might be necessary in order to be able to consider things aright.[54] This does not seem to help his position against Williams' complaint.)

Another way to think about this matter is to recall two remarks I made earlier about the absolute conception of the world. The absolute conception must explain both why the *Überwissenschaftleren* think what they do, and why some have thoughts different from the *Überwissenschaftleren*. The former explanation needs to show that their thoughts answer to the way things are, that they pass the reference-in-explanation test. The latter explanation needs to show why mere *Menschen* have the thoughts they do, correct or no. Now Williams is saying that the moral objectivist's claim that the immoral are in no position to deny that morality passes the reference-in-explanation test is suspect unless and until the moral objectivist can provide a good explanation for why the immoral person thinks differently. Then, and only then, is it reasonable to assert that the immoral person's conception of how to live is not just different, but also lacking. In order to ward off those who seek to show that morality does not pass the reference-in-explanation test, McDowell first needs to show that they are making some kind of mistake. The moral objectivist must discharge the second task in order for us to be convinced that she has adequately discharged the first.

[54] McDowell 1995.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

What has been shown? We began this section with the suggestion that someone could not reject morality across the board and yet grasp the virtuous person's moral views well enough to explain them, explain them as something other than an accurate conception of how things are. But we have seen that the moral objectivist cannot dodge questions about the proper explanation of moral views so easily. In order to convince us that the immoral cannot fully grasp the views of the fully virtuous, cannot grasp them well enough to explain them, the moral objectivist *himself* needs to explain why the immoral are deficient, why the immoral haven't grasped the truths grasped by the virtuous. The moral objectivist will convince us that the immoral cannot understand the moral views of the virtuous only if he first explains why the immoral have the views that they do. He must explain why those with 'incorrect' moral views fail to grasp objective moral truths, if he is to show that the reference-in-explanation test is an inappropriate gauge of the objectivity of morality. And so there is no avoiding an investigation into particular explanations of why people have the moral views that they do.

2.9 Conclusion

I have defended the thought that the objectivity of morality depends on its content entering into the explanations of facts whose existence is not so controversial, facts such as that we have views about what is right and wrong, good and bad. There

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

are several things to be said in defense of this thought: first, parsimony is an intellectual virtue; second, we want to avoid recognizing 'facts' that actually do no work; and third, insisting on such facts can be a way to demonize that which is merely different. The moral objectivist can honor parsimony, avoid fantasy, and treat his opponents fairly by showing that we need to call on morality in order to make sense of features of our lives that moral objectivist and anti-objectivist alike agree require explanation. The fairest way he can do that is by showing that the immoral are not in the best position to grasp truths about how to live. By doing so, the moral objectivist can *argue for* rather than merely reassert his position.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Chapter Three

The ins and outs of reasons

3.1 Introduction

We have seen that the moral objectivist, in order to deliver a satisfactory defense of his view, needs to show that the immoral are making some sort of error or mistake. If there are objective truths about how to live, the moral objectivist should explain why the immoral fail to grasp them. Otherwise, the moral objectivist is guilty of what Bernard Williams has called moralism.

It is clear that the moral objectivist ought not accuse the immoral person of error merely in virtue of the fact that the immoral person doesn't believe (or do) what others believe (or do). The moral objectivist owes us more than that. But there remains a further question about the nature of the moral objectivist's task. Does he need to explain why the immoral person fails to grasp moral truths to the satisfaction of immoral person herself? Does he need to convince *her* that she is mistaken? Or, rather, does he merely need to deliver good explanations for why the immoral person fails to grasp moral truths, where the standard for goodness is fairly independent of what the immoral person thinks about them?

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

One helpful way to reformulate these questions is to ask instead whether the immoral person can fail to believe (or to act) reasonably even if that person cannot see why she should so believe (or so act). Must a person be able to reach the conclusion that she should believe *p* (or that she should ϕ), if it is to be said truly that she has a reason to believe *p* (or that she has reason to ϕ)? That is, are all reasons internal? Or, rather, can it be the case that a person has a reason to believe *p* (or to ϕ) even if there is no way she could reach that conclusion herself? Can reasons be external?

If there are external reasons, then the mere fact that a person cannot see for herself that she has reason to believe moral truths (or to act morally) doesn't foreclose the possibility that she nevertheless has these reasons. She might be irrational even if she cannot see for herself that she is irrational. And so one might be able to explain that the immoral person is irrational for failing to grasp moral truths, even if the immoral person doesn't accept that explanation. This would open the door for the moral objectivist to argue that the immoral person fails to meet standards that truly govern her, even if he cannot convince her of this.

But *are* there external reasons? Must all of a person's reasons be potentially available to that person? Bernard Williams forcefully argues that all reasons for action are internal.[1] Williams thinks that an agent has no reason to act morally so

[1] What Williams thinks about reasons for belief is less clear. But it is more difficult to argue that there are external reasons for action, and so that is the task I will shoulder here.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

long as she has no desire that morality serves.[2] On his view, reasons get their grip on an agent only in virtue of her desires, motivations, plans and projects. To insist that an agent has reason to do something that in no way serves her desires is, he thinks, to say something false.

In this chapter, I will consider whether an agent can have reasons for action that bear little to no relation to what she wants to do. I hope to show, or make it seem a little more plausible, that *contra* Williams an agent's reasons for action can be quite independent from what she wants to do. We can accept Williams' demand that the moral objectivist provide some convincing explanation for why the immoral person fail to grasp truths about how to live, but, at the same time, resist his demand that this explanation need be one that the immoral person herself can accept. The moral objectivist can avoid moralism even if he cannot convince the immoral person of the error of her ways.

3.2 Explanation and Justification

I will first outline Williams' views of reasons for action, for I will argue for my view by showing what I believe to be the inadequacies of his. In developing his own theory of reasons for action, Williams distinguishes himself from two other

[2] Bernard Williams' views on internal and external reasons can primarily be found in Williams 1981, 101-113; Williams 1995a, 35-45; Williams 1995b; and Williams 1995c.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

characters, one of whom he calls *the external reasons theorist*, the other he calls *the sub-Humean*. The external reasons theorist maintains that the statement "Agnes has a reason to ϕ" does not imply anything at all about what Agnes wants to do.[3] On this view, Agnes' reasons are external to her set of desires. The sub-Humean, by contrast, thinks that Agnes has a reason to ϕ only if Agnes has some pre-existing desire that she believes will be served by her ϕ-ing.[4] That is, the sub-Humean thinks that all of Agnes' reasons are internal to her set of desires.

Williams thinks that neither theory is correct, each failing for different reasons. On his view, any adequate theory of reasons for action must meet two criteria:

> *(C1)* The theory should display how reasons can explain the actions that they are reasons for.
> *(C2)* The theory should display how reasons can justify the actions that they are reasons for.

Reasons have both an explanatory and a normative dimension, and any candidate theory of reasons is lacking if it fails to show how both are possible.[5]

---

[3] Williams does not specify who he considers to be an external reasons theorist, but the picture Williams paints bears some affinity to Kant, Tom Nagel, John McDowell, and, at moments, Philippa Foot. Some instead characterize Kant as an *internal* reasons theorist, but I believe this to be a mistake. Kant *was* an internalist, for he thought that seeing that a moral imperative applies to you necessarily motivates you to so act. But he was not an internal reasons theorist, because he did not think that our reasons were grounded in our psychology; instead Kant thought that our reasons were grounded in our rational nature, which he thought to be definitely not psychological.

[4] Williams does not say whether any philosopher is represented by the sub-Humean, but he does recognize that Hume's own views were more complex than those of the sub-Humean. Nonetheless the sub-Humean represents what many have taken Hume's views to be, and thus constitutes one of the standard accounts of practical reason that philosophers have considered.

[5] See Williams 1981, 102-103; Williams 1995a, 38-39; Williams 1995b, 191.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

One might question whether these are indeed proper criteria by which to evaluate a theory of reasons for action. Do reasons both explain and justify action? Couldn't it be the case that we have reasons to do things, even though there is no way for said reasons to explain what we do?

Or, rather, might it not be merely a verbal coincidence that we use the same word to refer to the things that explain action as well as the things that justify action? Is it a mistake to burden the reasons theorist with the task of capturing these two separate phenomena, much as it would be a mistake to burden someone who wanted to construct a theory of banks with the task of accounting for the nature of financial institutions as well as the nature of the slopes of land adjoining rivers?

I don't think that these charges stick; I think Williams is correct to demand that a theory of reasons for action meet both C1 and C2. Suppose that one and the same reason could not both explain and justify Agnes' ϕ-ing. This means that the thing that explains why she ϕed is necessarily different from the thing that makes her ϕ-ing the appropriate thing to do. Try as she might, Agnes cannot ϕ for the reason that makes her ϕ-ing appropriate. On this view, a person never acts for the reason that makes that action reasonable. And this seems just incorrect.

Better to think that, at least sometimes, Agnes ϕs for the reasons that favor ϕ-ing. A theory of reasons should show how this is possible. This is not to say that it will always be possible for the thing that justifies some course of action to explain her

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

doing it. In some cases these two dimensions of reasons may come apart. But insofar as the explanatory and the normative dimensions of reasons fail to track one another, a good theory of reasons for action will display why this is so, will display what about the agent or the circumstances made it the case that Agnes' reasons were explanatorily inefficacious.

Now Williams criticizes each of the two rival theories for failing to meet both C1 and C2. As we shall see in some detail later, Williams thinks that external reasons theory does not acknowledge the way in which reasons explain the actions that they are reasons for. If we divorce Agnes' reasons from what she wants, as the external reasons theorist recommends, then it is a mystery how her reasons are nevertheless supposed to explain her actions whenever she acts for a reason. External reasons theory fails to meet C1. Thus Williams embraces *internal reasons theory*, a theory that connects all of Agnes' reasons with what she wants.

But the particular version of internal reasons theory touted by the sub-Humean is wrong as well, for the sub-Humean fails to acknowledge the way in which reasons justify the actions that they are reasons for. To say that an action is rational is to say something good about it, and the sub-Humean view does not really show why this is so. Thus the sub-Humean view fails to meet C2.

But while Williams writes off external reasons theory as irredeemable, he thinks that he can iron out the wrinkles in the sub-Humean view by making three corrections to this otherwise cartoonish version of Hume's own thoughts. Williams'

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

first correction is to strengthen further the explanatory power of the sub-Humean model, enough to show that his version of internal reasons theory meets C1. The second and third corrections are to give internal reasons theory some normative bite, to enable internal reasons theory to meet C2. I will describe each of these corrections in some detail, and then examine whether Williams' considered view avoids the problems had by its rivals.

3.3 WILLIAMS' CORRECTIONS TO THE SUB-HUMEAN MODEL

The first modification to the sub-Humean model that Williams proposes is to ensure that his theory of reasons meets the explanatory criterion. Instead of insisting that all reasons for action imply the existence of some *desire* that the action serves, Williams takes a more liberal attitude to the topic. He maintains that the psychological states that can give rise to an internal reason are various, suggesting as candidates "dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be called, embodying commitments of the agent."[6] He refers to the set of all of the psychological states that can give rise to internal reasons as Agnes' subjective motivational set, or, for short, her S. On Williams' version of internal reasons theory, if Agnes has a reason to ϕ, then she has

[6] Williams 1981, 105. Williams specifically excludes beliefs and needs from his list of subjective motivations. The phrase "subjective motivation" itself is a technical term in Williams' writings, one whose meaning is not always clear. This will prove relevant later.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

some subjective motivation in her S that her ϕ-ing would serve. Whenever Agnes acts for a reason, we can explain her action by pointing to the relevant element in her S.

On Williams' view, this first correction strengthens the explanatory capacity of internal reasons theory. We might not always be able to explain Agnes' actions by pointing to some occurrent desire that her action serves. As Hume himself noted, we do not always feel the passions that explain our actions; sometimes they seem to operate calmly.[7] But in multiplying the types of subjective motivations that an action can serve, Williams thereby increases the chances that an internal reason can always explain Agnes' rational actions. The thought is that while our occurrent desires may not always explain our actions, our subjective motivations probably do.

The second correction has to do with the problem of false belief. Agnes can fail to act reasonably due to the fact that she is mistaken about some fact in the world; for example, if she does not know that she has just been poisoned, then she will not seek an antidote, even though she surely has reason to do so. When figuring out what reasons Agnes has, we are licensed to correct for her false beliefs. That is, we can help ourselves to the rationality governing belief when figuring out the rationality governing action.

---

[7] Hume 1978 [1739], II. 3. iii.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Williams' third and final correction to the sub-Humean model is related to the previous one. Not only are there rational standards governing belief, but there are also rational and irrational ways of thinking about what to do.[8] Consider the relation between some relevant element of S—call it the desire D—and the act of ϕ-ing. Internal reasons theory specifies that Agnes has a reason to ϕ just if D is *served* by ϕ-ing. The simplest way this relation can be instantiated is if ϕ-ing just is the object of D: Agnes might have a reason to ϕ simply because she wants to ϕ.

But Agnes might have an internal reason to ϕ even if ϕ-ing isn't the object of any of her subjective motivations, so long as she could ϕ as a result of rational deliberation from some D. For instance, if she rationally concludes that ϕ-ing would be the most efficient *means* to satisfying some D, then she may ϕ. Means-end deliberation is presumably rational deliberation, and the possibility of Agnes' taking the means to her end licenses the further claim that she has an internal reason to ϕ, even if she didn't have any subjective motivation to ϕ.

---

[8] Korsgaard 1986 notes that means-end explanations of action do not merely consist in the conjunction of a belief about the world and a desire. For Agnes might know the truth about the relevant causal relations between ϕ-ing and ψ-ing, and may want to ψ, but fail to be motivated by the consideration that ϕ-ing is a means to ψ-ing. The belief and desire alone may not be sufficient to motivate Agnes to take the means to her end. In such a case, we would say that Agnes is practically irrational, for while she knows the relevant causal relations, she fails to be moved by what is practically relevant. Means-end reasoning is a rational deliberative process, and it is not identical with the rational processes that characterize gaining knowledge about the causal relations between events in the world. Alternatively put, failing to have a belief about the world is not identical with failing to be motivated by certain considerations. This distinction noted by Korsgaard is one that Williams himself recognizes. Williams (1981, 104) notes that "the mere discovery that some course of action is the causal means to an end is not in itself a piece of practical reasoning", and he cites approvingly articles by Kolnai 1978 and Wiggins 1976, which he claims defend the distinction which Korsgaard later elaborated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

So much is generally allowed by followers of Hume, but Williams wants to be still more flexible about the matter. He recognizes that means-end reasoning is not the only form of practical reasoning; other likely candidates include

> thinking how the satisfaction of elements in S can be combined, e.g. by time-ordering; where there is some irresoluble conflict among the elements of S, considering which one attaches most weight to . . . , finding constitutive solutions, such as deciding what would make for an entertaining evening, granted that one wants entertainment.[9]

Williams even grants that the exercise of the imagination "about what it would be like if [some development] came about" can count as a form of rational deliberation: Agnes might discover through imagination that ϕ-ing serves her desires, something she would not discover had she been less imaginative. Williams thus leaves open the kinds of ways the service relation can be instantiated. There are a variety of ways a reason for action can serve an element of an agent's S, and this seems to make Williams' version of internal reasons theory truer to our ordinary picture of what kinds of reasons for action we have.[10]

---

[9] Williams 1981, 104.

[10] But it cannot be the case that any old thought process beginning from *D* and ending in Agnes' ϕing constitutes a rationally sound deliberative process. For instance, Agnes might have the following train of thought: "I want a haircut . . . my dog is kind of dirty . . . I wonder what's on TV right now . . . maybe I should call my mother . . . so I'll go get some ice cream", but such a train does not stay on the rails of rationality. (Candace Vogler drew my attention to the question of the status of these colorful flights of thought.) So not any sequence of thoughts constitutes sound deliberation. Later, we will be concerned with the grounds for distinguishing sequences of thoughts that *are* examples of sound deliberation from those sequences that are not. For now, however, we should simply note that Agnes may have many more reasons than motivations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Thus both by incorporating the rationality of belief into a theory of reasons for action and by distinguishing rational from irrational ways of practical thought, Williams extends the normative power of internal reasons theory well beyond that of the sub-Humean view. Williams' brand of internal reasons theory seems to capture both the normative dimension of reasons (C2) as well as their explanatory dimension (C1). Williams holds onto the basic Humean insight that action is grounded in our passions, and not just in reason, but is less rigid than most Humeans both about what can count as a passion, and about what can count as a reason.

3.4 Williams' argument against external reasons theory

Now Williams contrasts this emended version of internal reasons theory with the position favored by the external reasons theorist. On the external interpretation of reasons, if Agnes has a reason to ϕ, it is still an open question whether some element of Agnes' S would be served by her ϕ-ing. Making claims about the reasons Agnes has is one thing; making claims about the contents of her soul is something else. The external reasons theorist makes a sharp distinction between what it is rational for Agnes to do, and what she has a subjective motivation to do.

Williams, however, thinks that the external reasons theorist is mistaken to separate sharply these two items. He agrees with the external reasons theorist that reasons have a normative dimension, that they are used to highlight which courses of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

actions an agent is, in some sense, justified in doing. But reasons have an explanatory dimension as well, and if reasons are to explain why someone does what she does when she acts for a reason, it seems that they must display how she could be *motivated* to act in such a way. By failing to show how the action an agent has reason to perform hooks up with something she is dispositionally favorable towards, the external reasons theorist does not account for the ways in which reasons can explain the actions they justify. Or so Williams argues.

Thus if the external reasons theorist does not want to ignore the explanatory dimension of reasons, he must find some way to show how an agent's having an external reason could nevertheless explain her action, some way to meet C1. Various strategies present themselves here.

One possibility is that Agnes ϕs because she *believes* that she has a reason to ϕ. Such an explanation appears to be compatible with external reasons theory, for the psychological characteristic we cite to explain her action—a belief—is not the kind of element we would find in Agnes' S, for her S was said to include only subjective motivations such as desires, projects, loyalties, and the like. So her ϕ-ing may not serve any element of her S, but we could still explain her action by citing her beliefs about what reasons she has.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Williams, however, thinks that this suggestion fails to establish the sense in which there could be an external reason.[11] Suppose that Agnes does believe that she has a reason to ϕ. Now it is unclear just what the content of this belief is: *what* does Agnes believe when she believes that she has a reason for action? If the content of her belief is that she has some desire that would be served by her ϕ-ing, then this belief certainly explains her ϕ-ing, but it does so on terms favorable to the internal reasons theorist. For instance, if Agnes believes that she wants to go to the grocery store, and then she does go to the grocery store, it makes sense to suppose that Agnes indeed wanted to go to the grocery store, and that this want helps to explain her action.

Alternatively, the content of the belief may not itself *cite* an element of Agnes' S, but, given the nature of the belief and the fact that Agnes ϕs, we may nevertheless appropriately presume that Agnes acted for a reason. For instance, the external reasons theorist might hold that Agnes ϕs, not because she is subjectively motivated to ϕ, as the internal reasons theorist thinks, but rather because she thinks that it is morally right for her to ϕ. The fact that it is morally right to ϕ seems to be Agnes' reason for acting. It seems that this could be an appropriate explanation of Agnes' action.

[11] Williams 1981, 107.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But this explanation doesn't appear to tell against the version of internal reasons theory Williams defends, for, if the thought that ϕ-ing is morally right indeed explains her action, then it seems very likely that it does so also because Agnes *wants* to do what is morally right. The belief and the want (or other subjective motivation) work in tandem to explain the action. Surely the external reasons theorist will not insist that Agnes does *not* want to do what is morally right. So the link with Agnes' S appears to be preserved, and Williams' version of internal reasons theory seems to escape from this objection unscathed.[12] This alleged explanation of Agnes' action that the external reasons theorist might be tempted to offer turns out to be inadequate for his purposes.

But if neither of these options is the correct elucidation of the content of Agnes' belief, then we are without a full explanation of her action. Her belief is supposed to explain her action, but if we don't really understand the content of what she believes, we haven't yet an satisfactory explanation of what she does. All satisfactory attempts to flesh out the content of what Agnes believes also make it plausible to think that she has some subjective motivation that her action serves. The only way the external reasons theorist can meet C1 is by adopting a position actually favorable to the *internal* reasons theorist.

---

[12] Williams 1995a, 37-39.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Williams grants that the external reasons theorist could make her position plausible by showing how Agnes could *come to believe* that she had a reason, come to believe it in a way other than by deliberating from some element in her S. If Agnes gains the belief that she has a reason to ϕ, and she *thereby* acquires a new subjective motivation that her ϕ-ing would serve, external reasons theory would be true. But she has to acquire the belief in the right sort of way; it does the external reasons theorist no good if Agnes comes to believe she has a reason to ϕ as a result of being persuaded by sophistry. *That* wouldn't show that Agnes had an external reason all along, a reason she has just discovered. Rather, the external reasons theorist would be vindicated only if Agnes comes to believe that she has a reason to ϕ because she is "considering the matter aright." Only by reaching the belief in a reliable, rational way does Agnes' belief-acquisition give the external reasons theorist the kind of result necessary to vindicate his position.[13]

But Williams doubts that the external reasons theorist can actually show this. Williams can see no way for Agnes to suitably gain this belief except by deliberating *from* some pre-existing subjective motivation:

> For, *ex hypothesi*, there is no motivation for the agent to deliberate from, to reach this new motivation. Given the agent's earlier existing motivations, and this new motivation, what has to hold for external reason statements to

[13] Williams 1981, 108-109. Note that Williams' strategy here is to argue that it does us no good to deliver a satisfactory account of the explanatory dimension of reasons, if we thereby forfeit our ability to deliver a satisfactory account of the normative dimension of reasons. This is why Agnes has to acquire her belief in a reliable way. Later on, I will adopt a similar strategy in my argument against Williams' view.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> be true, on this line of interpretation, is that the new motivation could be in some way rationally arrived at, granted the earlier motivations. Yet at the same time it must not bear to the earlier motivations the kind of rational relation which we considered in the earlier discussion of deliberation – for in that case an internal reason statement would have been true in the first place. I see no reason to suppose that these conditions could possibly be met.[14]

All rational deliberation must start from something that is wanted; practical reason does not operate in the void. It does not appear that the external reasons theorist can convincingly display how Agnes could come to believe that she has a reason to ϕ, unless it also has been true all along that she has had some subjective motivation that this reason serves; that is, unless she has an *internal* reason to ϕ.

So if reasons are to explain action, and if only internal reasons theory can give an actual account of how reasons—or even belief in reasons—explain action, then to insist on external reasons theory anyway is mere bluff.[15] Statements that explain Agnes' action are, in part, statements about her psychology, and external reasons theory has not shown that it has the resources to say enough about Agnes' state of mind. By contrast, Williams' notion of an agent's S is liberal enough such that it seems we can always link up Agnes' actions with some element in her S. The resulting version of internal reasons theory thereby provides a framework within which reasons can explain action. Thus only internal reasons theory, Williams argues, does justice to C1.

---

[14] Williams 1981, 109.

[15] Williams 1981, 111.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.5 THE VACUITY OF WILLIAMS' ARGUMENT

Matters take on a different tone, however, if we examine more closely the nature of the elements in Agnes' S. Williams' description of the various elements that make up her S has been very vague. I don't simply mean that by leaving open-ended the kinds of psychological characteristics that can be elements of her S, Williams' position is hard to understand fully, and thus, hard to evaluate, though that may be true too.[16] Rather, we haven't much indication yet what a subjective motivation even is, how it is to be thought of, and what the conditions are for describing someone as having one.

For consider the following. As we have seen, the external reasons theorist maintains that Agnes can have a reason to ϕ, and that her acknowledgment of this reason can play a role in the explanation of her intentionally ϕ-ing, despite the fact that Agnes might have had no subjective motivation to ϕ, nor any subjective motivation that would be served by ϕ-ing. Williams rightly challenges the external reasons theorist to spell out how such explanations are to work. But a more fundamental challenge to the external reasons theorist is to question the legitimacy of his descriptions of Agnes' psychological states. What grounds has he for describing

[16] For criticisms of this nature, see Cohon 1986.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Agnes' soul in such a way? Surely not on the basis of her actions, for the fact that Agnes ϕs intentionally seems to be powerful evidence that she wanted to ϕ, or that she wanted something else that she could attain by ϕ-ing. The external reasons theorist needs to tell some story about how he can reliably specify the content of Agnes' S, then, if he is going to discount what consideration of Agnes' actions suggests. This he hasn't done yet.

This suggests that there is some kind of *conceptual* or *grammatical* connection between ϕ-ing intentionally and wanting to ϕ (or wanting something that ϕ-ing is a step towards). If Agnes ϕs intentionally, she *ipso facto* has a want that her ϕ-ing serves.[17] But it is also true that if Agnes ϕs for a reason, she ϕs intentionally—this seems obvious. But then it follows that whenever Agnes ϕs for a reason, she has a want that her ϕ-ing serves. Internal reasons theory, then, appears to be vindicated by this conceptual connection between wanting and intentional action. Whenever Agnes acts for a reason, it follows that she does something that serves one of her wants.

Now, usually in philosophical argument, if you can show that your thesis is a conceptual truth, or that it is true by definition, you have won your argument, for you have shown that anyone who opposes your view is just confused. But here things are

[17] Of course, there is a gap between 1) acknowledging that the intention is powerful evidence for the existence of the want, and 2) claiming that the existence of the intention *ipso facto* implies the existence of the want. We would need a full account of the concept of intention, such as Anscombe 1957 to license this transition. But even if we are stuck with only the first of these positions, I still think we could mount an attack on the kind of external reasons theory that Williams discusses.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

different. Williams thinks that a theory of reasons needs to be psychologically realistic, and furthermore that internal reasons theory makes room for psychological explanations of actions, while external reasons theory does not.[18] Presumably, psychological explanations are substantive; they have content. So if reasons are to explain the actions that they are reasons for, they should do so in ways such that it also makes some sense to deny them.

But we have seen that the explanations that internal reasons theory affords aren't substantive at all. Instead we have only identified a sense of the word 'want' such that whenever Agnes does something for a reason, it can be said that she wanted to do it (or wanted something that her action was a step towards). Williams' version of internal reasons theory is true just by definition; the explanations that it offers are trivially true. Thus explaining Agnes' ϕ-ing by automatically ascribing to her a subjective motivation isn't to give a substantive explanation of what she does. And so these explanations aren't *psychological* explanations. That is, all of Agnes' intentional actions speak to something she wants because of the way in which the concepts of wanting and of intentional action depend upon one another, not because Williams' brand of internal reasons theory has been confirmed by psychological insights.

This should have been obvious earlier when we were looking at Williams' argument against the external reasons theorist. Whenever the external reasons

---

[18] Williams 1995b, 191.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

theorist tried to explain Agnes' action by pointing toward one of her beliefs, Williams announced that Agnes also must have had some desire (or subjective motivation) that helped to explain what she did. But this insight did not fall out of an extensive psychological study of Agnes' life; rather, Williams describes Agnes as having a subjective motivation to ϕ simply in virtue of the fact that she so acted. If Agnes intentionally moves, she has a motive. To say that Agnes has a desire that her ϕ-ing serves is to say no more (and maybe to say much less) than that she has a reason to ϕ.[19]

Thus Williams has not shown that his view is actually superior to external reasons theory on explanatory grounds. Everything that Williams says about the close connection between rational action and subjective motivation could be true, and yet we have no grounds for thinking that subjective motivations substantively explain rational action, no grounds for thinking that Williams has met C1. And so we haven't any reason yet for rejecting external reasons theory in favor of its rivals, not according to the criteria for a satisfactory theory of reasons for action that Williams himself endorses.[20]

---

[19] Some philosophers of mind, however, are happy to acknowledge that there are conceptual connections between mental causes and effects; they do not view this as a problem for the explanatory capacity of mental states. To this I respond that it all depends on the nature of the conceptual connection. If the explanandum alone conceptually implies the supposed explanans, as Williams' view seems to suggest, then the putative explanation is indeed vacuous. But we can have a substantive explanation if the explanans is conceptually related to many potential and actual explananda, no one of which is sufficient to license the inference that the explanans exists.

[20] The classic sources informing the view I present in this section, different though they may be, are Anscombe 1957 and Nagel 1970. See also McDowell 1995.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

3.6 SEEKING NORMATIVITY

Let us suppose, however, that my argument in the previous section fails for some reason that I haven't foreseen, and that Williams' version of internal reasons theory is *indeed* superior to external reasons theory on explanatory grounds, that it does meet C1. The focus of attention then shifts to whether Williams' theory adequately displays the normative dimension of reasons (C2). Otherwise, the external reasons theorist can still reasonably maintain his view on the grounds that while his theory makes sense of the common thought that irrational action is in some sense bad action, internal reasons theory does not.

Williams attempts to give internal reasons theory some normative bite through the second and third corrections to the sub-Humean model. On his view, Agnes can fail to act reasonably either because she has false beliefs, or because she does not deliberate soundly, e.g. she does not take the means to her end, she is unimaginative, she fails to find constitutive solutions, and so on. We can criticize Agnes' action, then, in virtue of these flaws in her thinking. Thus it seems that Williams has crafted an internal theory of reasons that meets C2.

However, I will argue in this section that Williams' final position has the following problem: in his attempts to make internal reasons theory account for the *normative* dimension of reasons (C2), Williams loses any supposed advantage he had

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

over external reasons theory in accounting for the *explanatory* dimension of reasons (C1). That is, Williams forfeits any explanatory superiority his view had over his rival when he tries to do what his rival does well.

Notice that on Williams' final view, as on the view of the external reasons theorist, Agnes can have a reason for action that cannot explain what she does. For instance, as we saw earlier, if Agnes does not know that she has just been poisoned, then she will not take an antidote, even though she surely has a reason to do so. So long as she remains ignorant about the fact that she has been poisoned, her reason to take the antidote will not explain what she does. Of course, if Agnes *were* well-informed, the reason she now has might explain what she *would* do. Thus even here the reason is not entirely divorced from its explanatory role; rather, it potentially explains the action of Well-informed Agnes, not Actual Agnes. So even though Actual Agnes' reason to take the antidote still presumably serves some element in her S, it cannot explain what she does.

Similarly, Williams allows that Agnes can have a reason to ϕ, even if it cannot explain what she does, this because Agnes cannot deliberate soundly. For instance, Agnes might intend to ψ, and believe that ϕ-ing is the only means of ψ-ing, and yet still not ϕ. (Let us presume that ϕ-ing doesn't interfere with any of her other ends, and is not particularly grueling nor immoral nor otherwise *un*reasonable.) Since means-end reasoning is presumably one way of deliberating well, Agnes' failure to take the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

means to her end suggests, not that Agnes does not have a reason to ϕ, but that she is irrational.

So even on Williams' view, Agnes can have a reason that cannot explain what she does. Reasons explain only what an agent *would* do were she to meet certain criteria—such as being well-informed, or being able to take the means to her end—not what she in fact does. The perspective from which reasons explain action has shifted.

Now if Williams can alter the explanatory requirement in order to accommodate the normative dimension of reasons, perhaps the external reasons theorist also can find room for altering the explanatory requirement in ways that promote his position. To see whether the external reasons theorist can do this in ways that won't seem *ad hoc*, we will need to consider what grounds Williams has for altering the perspective from which reasons explain action. Why does Actual Agnes have a reason to ϕ, if such a reason could explain only what Well-informed Means-taking Agnes would be motivated to do? Williams answers

> The internalist proposal sticks with its Humean origins to the extent of making corrections of fact and reasoning part of the notion of 'a sound deliberative route to this act' . . . . The grounds for making this general point about fact and reasoning . . . are quite simple: any rationally deliberative agent has in his S a general interest in being factually and rationally correctly informed. There could be a case of somebody who had an overwhelming need to be deceived; and if his relations to reality were so poorly negotiated that he actually needed to believe what was false, then perhaps he would have reason to acquire false beliefs—in that particular respect. The basic point, however, is that on the internalist view there is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> already a reason for writing, in general, the requirements of correct information and reasoning into the notion of a sound deliberative route . . . [21]

We can detect two competing threads in his response, threads which do not go well together. The first is that "any rationally deliberative agent has in his S a general interest in being factually and rationally correctly informed." The idea here is that Agnes has a *desire* (or project, disposition, or other subjective motivation) to be well-informed, and thus reasons can be fairly attributed to her in light of what could explain her actions if she were indeed well-informed. The grounds for idealizing her condition in this way rests on what *Agnes* wants, and since she wants to be well-informed, her reasons issue from the perspective of what she takes to be an ideal version of herself, viz. Well-informed Agnes. We can say that Agnes has a reason to ϕ, because she would ϕ if she were the way *she* wants to be. Let us call this option O1.[22]

The second thread in Williams' response does not find the requirement to correct for false beliefs and irrationality *within* Agnes' S, but instead writes it "into the notion of a sound deliberative route" *from* her S to her ϕ-ing. "Making corrections of fact and reasoning [is] part of *the notion* of 'a sound deliberative route to this act'

---

[21] Williams 1995, 37. Notice that Williams doesn't deny that the person with the need to be self-deceived does not *also* have a reason to acquire true beliefs. Also, elsewhere Williams denies that an agent's reasons are internal to her needs, so his reliance on needs in this passage is puzzling. See Williams 1981, 105.

[22] Williams (1995c) unambiguously embraces this first approach: "What A has reason to do is thus conceived of as a projection of A's psychology, with improvements *to the extent that A is interested in* such improvements . . . " [*Emphasis* mine]. For an intricate and well-developed instance of this type of view, see Rosati 1996.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"[*Emphasis* mine].[23] Here is seems that Actual Agnes has reason to do what she would do if she were well-informed and took the means to her ends, not because *Agnes* happens to want to be that way, but because it is just part of deliberating well to do so rationally in light of the truth. On this view, the notion of sound deliberation isn't built out of what the particular agent in question counts as sound deliberation. Rather, the qualities that characterize sound deliberation—qualities such as having true beliefs and taking the means to your end—somewhat transcend the features of any individual's psyche. Actual Agnes has reason to do what Well-informed Means-taking Agnes would do, because being well-informed and taking the means to one's ends are necessary for deliberating well. Let us call this option O2.

These two ways of thinking about the gulf between Agnes' actual subjective motivations and her reasons signal different possibilities for the moral objectivist. O1—the option that bridges the gulf by building out from Agnes' own S—does little for him. Agnes may have reasons that cannot explain what she does, but this is only because she wants to be the kind of person who would be motivated by such reasons. These reasons are still 100% internal, ascribed after her S has been subjected to a thoroughly *internal* critique.

O2—the option that bridges the gulf by applying an independent notion of sound deliberation to Agnes' S—is more promising. Since the notion of a sound

---

[23] Williams 1995a, 37.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

deliberative route is not built out of Agnes' S, it is a very open question what Agnes has reason to do. If a sound deliberator is well-informed, skilled at technical means-end reasoning, and so on, then Agnes may have many more reasons to do things than she is now motivated to do. And who knows what other qualities constitute a sound deliberator? On this approach, the answer to this question is not simply up to Agnes.

So, which of these two ways of understanding the grounds for altering the explanatory requirement is correct? Are the qualities that constitute sound deliberation essentially peculiar to each agent, wholly built out from her own psychological states? Or does *every* agent have reason to do the things that she would be motivated to do were she well-informed and so forth?[24]

Well, suppose Agnes especially prizes freespiritedness. Agnes is under the impression that she just plain thinks too much, she dawdles, and she would act more rationally if she would trust her instincts more. (Perhaps she thinks that she is paralyzed by fear of the unknown.) Suppose, however, that Freespirited Agnes would not prize her free spirit in the way that Actual Agnes does. That is, suppose that if Actual Agnes suddenly became as freespirited as she now thinks she should be, she would end up thinking—to her surprise—that her devil-may-care attitude does not generally enable her to grasp better the reasons she has. Freespirited Agnes would

[24] I now set aside the exegetical question concerning which of the two options is Williams' most considered view. He definitely leans toward the first interpretation, but there is good textual evidence for the second option as well.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

actually regret the change, and would prefer to become patient again, to stop doing things hastily.

If the quality Actual Agnes prizes is not one she would want if she were to acquire it, then it seems to be a mistake to think that the perspective that she currently thinks of as ideal does in fact equip her to act upon her reasons. In such a case, Agnes seems to have gotten things wrong. Of course, the mere possibility of this kind of conflict does not show that it is in fact unreasonable for her to act in a freespirited manner, but it does reveal something about the basis for why it would or would not be unreasonable. The fact that we *might* regret having the qualities that we now believe constitute sound deliberation suggests that we can be mistaken about such things.

But perhaps philosophers can arrange things such that Agnes regrets acquiring the one quality that she really needs in order to act rationally. For instance, couldn't it be the case that while Actual Agnes thinks she would act more reasonably if she were more knowledgeable, Well-informed Agnes would prefer not to be well-informed, regretting the increase in her knowledge, *without* this showing that Actual Agnes was wrong to begin with? Maybe the qualities people need to deliberate well are such that those who have them think that they shouldn't. If so, Actual Agnes wouldn't necessarily be wrong about what qualities constitute sound deliberation.

But if Actual Agnes is correct to think that being well-informed would enable her to act upon her reasons, and if Well-informed Agnes disagrees with Actual Agnes, then Well-informed Agnes is wrong about what constitutes sound practical thinking.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

If an agent comes to have a different opinion about what she needs in order to deliberate well, she will not be right both times. Agents can be wrong about what constitutes sound deliberation.

If we deny this conclusion, then we are led to unattractive implications. An agent who changes her view about what constitutes sound practical thinking will *herself* often think that she earlier had been mistaken. For instance, if an agent comes to think that she needs to be more well-informed in order to act rationally, then she generally will also think that she had been mistaken to have thought that she didn't need to be well-informed. But if we deny that agents can be wrong about what constitutes sound deliberation, we would have to conclude that agents are often wrong when they think about their own earlier thoughts in this way. We would be forced to convict them of error whenever they believe that their views about practical thought have improved. Rather than embracing that unattractive implication, it seems preferable to acknowledge that agents can be wrong about what constitutes sound deliberation, which thereby makes room for acknowledging that agents' views about their earlier thoughts are sometimes right. That is, it is more charitable to take the position that agents are sometimes right about both matters.

Thus it seems that the qualities that constitute sound deliberation cannot plausibly be built out from each agent's psychological states.[25] The fact that an agent

[25] For a similar conclusion, see Lear 1984, especially 154.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

thinks she would be a better deliberator under different conditions does not guarantee that she *actually* would be a better deliberator under those conditions. And so the mere fact that an agent thinks she would deliberate better if she had a certain quality does not imply that her actions would be rationally justified if she were to acquire that quality. That is, her actions might be even less rationally justified if she were to become the way she thinks she should be.[26] So a theory of reasons cannot capture the normative dimension of reasons simply by relying upon particular agents' conceptions of practical rationality. Understanding the grounds for altering the explanatory requirement of reasons as the first approach recommends makes it impossible also to understand the normative requirement of reasons. Following O1 will not enable us to attain C2. Rather, a satisfactory theory of reasons must display external constraints on the notion of sound deliberation.

Since it is not up to each individual agent to decide what constitutes sound practical thinking, O2 seems more appropriate. This means that the notion of a sound deliberative route can be extended in ways that Agnes might not acknowledge as sound. She may have reason to do the things that she would do if she were more, say, imaginative, even if she doesn't now think she should be more imaginative.

But in any case, it seems that the notion of a sound deliberative route will still always *start* from something Agnes wants; even the most controversial way of

---

[26] An examination of Aristotle's uneven catalogue of the virtues would have reminded us of this. See *Nicomachean Ethics*, IV. iii on the great-souled man. What Aristotle took to be a virtue strikes most contemporary readers as a vice.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

thinking about what to do appears to have its origins in Agnes' S. Reasons, on this second approach, turn out to be neither wholly internal nor wholly external. They originate in what the agent wants, and so are technically internal, but they can be extended in all sorts of ways the agent may not acknowledge, and so are of a variety that the would-be external reasons theorist might be able to live with.

In this section we have been concerned with how Williams could legitimately alter the explanatory requirement in order to capture C2, the normative dimension of reasons. On the second interpretation of how he can do this (O2), reasons explain not what Agnes does, but what she would do if she deliberated well, where the standards of sound deliberation are *not* necessarily derived from what Agnes takes them to be. This hybrid theory of reasons seems to account both for the explanatory dimension of reasons—insofar as it shows how reasons could explain what some idealized version of Agnes would do—and for the normative dimension of reasons—insofar as it displays how acting on her reasons involves avoiding various kinds of error.

3.7 Mixing Minds and the Limits of Philosophical Surgery

So what is wrong with this second approach to altering the perspective from which reasons explain action? Wouldn't this hybrid theory enable Williams both to incorporate a normative element into internal reasons theory and to hang onto his (questionable) explanatory advantage?

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I want to argue that the resulting hybrid theory still bears significant problems. First, consider what qualities actually form the notion of a sound deliberative route. Williams is deliberately noncommittal in his answer to this question, but we do know that he thinks that a sound deliberator has true beliefs, takes the means to her ends, is imaginative, finds constitutive solutions to deliberative problems, and has many other qualities. Let us call the conjunction of these qualities Q.[27]

So the hybrid view maintains that Agnes has a reason to ϕ only if such a reason could explain the actions of an agent bearing on the one hand Agnes' S, and on the other hand the qualities Q. We paste the cognitive/deliberative excellences of the ideally rational agent onto the conative faculties of Agnes, and then consider what reasons could explain the actions of this interesting mongrel. Such an agent would begin with the same subjective motivations that Agnes actually has, but would be able to deliberate *from* these motivations as well as anyone. The qualities Q constitute the normative standards of practical rationality; Agnes' S provides the oomph that explains action.

Two related problems with this proposal come to mind. One falls under the rubric Davidson and others have called *the holism of the mental realm.*[28] The very content of a mental attitude is modified, mediated, and even constituted by the other

---

[27] Q could also include having undergone relevant educative life-experiences, if our theory of practical rationality made room from these. But I am getting ahead of myself.

[28] Donald Davidson, "Mental Events" in Davidson 1980, 207-225.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

aspects of the mind in which that attitude is located. Specifically, the content of an agent's conative attitudes depends heavily upon what she believes, what she hopes for, what she fears, and many other things about her state of mind. And the direction of dependency runs the other way as well.

This implies that we may be unable to pry Agnes' subjective motivational set away from the rest of her personality, pry the qualities Q of the ideal deliberator away from the rest of *her* personality, and then slap them together to get a coherent and unified agent that provides the perspective from which reasons explain action. The contents of the mind may be too sensitive to one another to survive this sort of violent relocation. We cannot successfully transplant mental qualities in our arguments like we can hearts and livers in real life, for the host environment might reject (or radically alter) the transplanted newcomer. There would probably be too much loss and distortion; only the crudest faculty psychologist would think otherwise. External standards of sound deliberation are not designed to cope with agents whose motivations are opposed to or radically different from those embodied in the standard. Thus the whole idea that such a hybrid agent even could exist is somewhat dubious.

This is easy to overlook when focusing upon certain kinds of cases; for example, the otherwise normal agent who is simply ignorant of the fact that she has been poisoned. Here it seems that all we need to do to show how she might act upon her reasons is to make her aware of this single fact. But if we focus instead upon the case of the callous unfeeling soul, for example, the person who just cannot be brought

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to care for her friend, it seems more doubtful that we will be able to discern her reasons for action by asking what a rational deliberator with those and only those desires would do.[29] For it seems likely that rational deliberators would not have *her* set of desires; their (whole) state of mind is probably quite different.[30] Of course, even the best of agents will encounter times and circumstances when their affection for others is flagging, and they will have resources and strategies to cope with meeting their responsibilities when their capacity for sympathy is muted. But they probably will not be of much use as an analytic device for gauging the reasons had by the poor creature whose affection for her friend is completely extinguished. We may be unable to display her reasons simply by pasting some but not all of the psychological features of an ideal agent upon her subjective motivations. And since this hybrid theory has not articulated a coherent perspective from which reasons explain action, it has not met C1.

A second and related problem has to do not so much with whether Agnes' S and an ideal deliberator's Q *can* be contained in the same mind, but with whether the resulting hybrid agent, if it *did* hold together, would even accurately display Agnes' reasons. Grounds for doubt arise when we consider whether agents have reasons to

---

[29] This is basically an example Williams uses. See Williams 1995a, 39.

[30] At the very least, a rational deliberator would not get herself into a position to *be* friends with someone she didn't care about. Of course, on a sufficiently broad conception of practical reason, a rational deliberator would care about anyone who was her friend in virtue of the fact that they are friends. But this is not the place to defend particular conceptions of practical reason.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

act in ways that would make them better deliberators. For example, suppose that Williams is correct to think that a sound deliberator does not commit logical fallacies in her thinking. Now if Agnes herself is a careless thinker, we might think that she has reason to act in ways that would help her avoid such mistakes; for instance, she might have reason to take a course in elementary logic. But the hybrid agent we have been discussing does not have any reason to take such a course, for she *already* avoids fallacies in her reasoning. The hybrid agent does not have reason to improve her deliberative capacities in the way that Agnes seems to have, and so it appears that we will misgauge Agnes' reasons for actions if we focus simply upon what would explain the actions of the hybrid agent. Agnes seems to have reason to study logic, even though this reason could not explain what the hybrid agent would do.[31] So even if we can make sense of the hybrid agent that Williams needs in order to solve the problem about normativity, we still will not have found the perspective from which reasons explain action. Again, we will have failed to meet C1.

Thus Williams fails to display how reasons both explain and justify action. The first interpretation of Williams' theory that we considered (O1) does not adequately capture the sense in which reasons justify the actions that they are reasons for. Since agents can be wrong about constitutes sound deliberation, this interpretation fails to meet C2. But O2 doesn't work either. We do not capture the

[31] Williams himself makes a similar point when arguing against the external reasons theorist's proposal that a sound deliberator already has the ethical virtues. He appears not to see that the same argument can be applied to the cognitive virtues as well. See Williams 1995b, 190.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

perspective from which reasons explain action by importing the deliberative virtues of the ideally rational agent into the mind of the ill-motivated one. So this second interpretation fails to meet C1. Thus Williams' theory, no matter which way we take it, is unsatisfactory according to his own criteria.

3.8 Implications

If we need not embrace internal reasons theory, then we may say, without bluffing or table-thumping, that a person can fail to grasp her reasons for action, even if she herself does not think that she is in a poor position to grasp them. An agent may fail to have the qualities necessary for sound deliberation without recognizing that this is so.

This opens up the possibility that the immoral person fails to act reasonably, even if she does not want anything that moral activity would serve. Morality's grip on her may not depend on its hooking up with anything she wants. The deliberative qualities that characterize the practically rational person may be incompatible with a life of immorality, even if the immoral person thinks otherwise. And so in acting immorally she may thereby fail to meet normative standards that nevertheless have authority over her. The immoral person might not be merely different from the moral person; she might be in error.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

To see whether this is so, we need to determine what qualities actually constitute sound deliberation, independently of what any particular agent thinks about the matter. Surely deliberators should be well-informed, but what other qualities characterize the rational agent? And will an agent with these qualities be committed to morality?

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER FOUR

Are the immoral ignorant?

4.1 LACKING INFORMATION

1. I have argued that the defender of moral objectivity must convincingly explain why some have incorrect views about what to do, while others see things aright. The claim that morality is objective depends upon whether the upright grasp truths about how to live, truths that the immoral do not detect. Claims that our own views about how to live are explained by their content are suspect unless and until we can also explain why others' views about how to live are *not* explained by their content, explain why others are in a poor position to grasp practical truth. Otherwise, we fall prey to what Bernard Williams has called moralism.

But can we show that the immoral have cognitive deficiencies? Why do the immoral fail to live morally? What makes for a gap between a correct view about how to live, and a person's actual view about how to live?

One of the most common ways to explain that someone has failed to grasp some truth p is to show that they have failed to grasp some other truth q, something they should know if they know whether p. The truths in question may be related to one another in various ways.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Perhaps *p* and *q* are inferentially related to one another, such that you could probably infer *q* if you knew *p*. Suppose that you are talking with someone else about the intricacies of the balk rule, and in the course of conversation it emerges that your interlocutor does not know that in baseball the runners can steal bases. Discovering that he does not understand this basic fact about the game gives us good reason to think that his thoughts about the details of the balk rule are unsound. Since the whole point of the balk rule presupposes that base runners try to steal bases, not understanding this latter bit of information implies that one has a poor grasp on the former. We can confidently and 'unmoralistically' reject what he has to say about the balk rule. We can see why he is in error.

Alternatively, perhaps *p* and *q* are related to one another such that if you were in a good position to know one, you would be in a good position to know the other. Suppose at a cocktail party you meet an older man who tells you that as a youth he was a student of none other than Ludwig Wittgenstein. He appears sincere, but you have heard that your interlocutor is susceptible to delusions of various sorts. Wondering whether he is now telling the truth, you steer the conversation in ways that reveal whether he has knowledge of the details of the life of his supposed teacher. As you ask him various questions about Wittgenstein, it becomes clear that your interlocutor does not believe some basic truths about Wittgenstein ("Austrian accent? Why, Wittgenstein could have passed for King George himself!") Anyone who met Wittgenstein would know that he spoke with a thick accent. We can confidently

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conclude that our dear old man does not know whether he had Wittgenstein as a teacher, because we have discovered that he doesn't know some things he would were that true. We often can justify our claims that someone does not know whether *p* by demonstrating that they are ignorant about something else that they would be in a good position to know if they knew whether *p*.[1]

So one strategy for showing that the immoral lack knowledge about what to do, then, is to show that they are ignorant about some other, 'nonmoral' matter that one would likely know if one's practical views were correct. According to this hypothesis, the immoral have false beliefs about matters true beliefs about which are strongly correlated with having sound views about how to live. Citing these false beliefs would constitute demonstrating that they are not in a position to grasp practical truth.

So perhaps the wicked wrongfully harm other people because they have false beliefs about the others' modes of life. Perhaps they lie because they falsely think that they are likely to get away with it. Perhaps they are ungrateful because they do not have a sound conception of what kind of effort was exerted on their behalf. Perhaps they are greedy because they falsely believe that money will satisfy them. Perhaps they attack their neighbors because they falsely believe that otherwise their neighbors will attack them. Perhaps they pressure one another to commit suicide because they

[1] This is also a common way for lawyers to discredit witnesses: at least it is common in the movies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

falsely believe that the extraterrestrials will snatch their souls as they ascend past the nearby comet. In each of these cases, if we can establish that the immoral have these beliefs, and that what they believe is false, and that having these false beliefs makes it unlikely that their practical views are correct, then perhaps we can satisfactorily show that their immorality is really a mistake.[2] It seems that then we would have an adequate theory of error.

No doubt, much wrongdoing stems from ignorance. Often I am led to act wrongly if I have a false view about those whose interests I interfere with or ignore. I am led to act wrongly also by having a rosy view of myself, my kin, friends, and nation, our collective past, and our intentions for the future. Ignorance of the multifarious connections between action and consequence likewise contributes to all kinds of malfeasance. Many instances of immorality can be chalked up either to false beliefs about others, ourselves, and the nature of the world.

But I think that, for better or worse, it is overly optimistic to think that we can explain the bulk of immorality in this way. In some cases we can surely determine that someone's practical views are unreliable by noting what else she believes. But I seriously doubt that most differences in practical views can be strongly correlated with ignorance about (other) factual matters. The moral life, whether a result of

[2] Sometimes ignorance excuses a person's wrongdoing; sometimes it does not. An exhaustive investigation into the explanation of wrongdoing would have to sort out these cases, but doing so here would take me too far afield.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

knowledge or not, is far too complex a thing to be understood as a function of our views about the nonmoral facts.

I haven't any explicit argument for my suspicion that 'nonmoral' knowledge is not sufficient for moral action. I am guessing, however, that most readers will, upon due reflection, agree with me. They will realize that it is a caricature of the best of the Enlightenment—and an ignoring of the worst of our century—to hold that the removal of false beliefs will usher in an era of justice.[3] But the most effective way for me to sway the unconvinced reader of this will be to point toward *other* ways to explain immorality, ways which have little to do with possessing information. That way we can see that having good information isn't enough to secure moral action. It will be clearer that one can be well-informed and yet act immorally. And so the hope to explain incorrect moral views in terms of ignorance about the nonmoral facts can at best be only partially fulfilled.

4.2 Bad philosophy

How else, then, might we link immorality with error? It has become more and more common for philosophers to focus upon the ways in which our judgments about

[3] On the other hand, acquiring correct information is often sufficient for a person to begin to respond to the injustice that she herself suffers. I don't think, however, that acquiring correct information is generally sufficient for a person to stop acting unjustly, or otherwise immorally.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the rightness and wrongness of particular actions should be responsive to the demands of the more general *theories* we accept.[4]

On this view, when making particular moral judgments, we are to consider their plausibility in the light of the more general moral principles we accept, with an eye toward reconciling the recommendations of the one with the other. That is, we might have good reasons for holding a particular moral theory (e.g. utilitarianism) as well as good reasons for making certain particular moral judgments (e.g. the amount of pleasure experienced by rapists is irrelevant in assessing the rightness/wrongness of their actions). If we find that our moral theory and our particular judgments do not fit one another—if our views are incoherent—then we are to reflect on the reasons why the principles seem sound and the reasons why the particular judgments seem sound, back and forth making adjustments in one when it seems appropriate in the light of the other, until we have arrived at an overall view that is coherent. It is possible that we will make the same particular moral judgments as we did before, but it is much more likely that reflecting on the general moral principles we accept will cause us to make *some* adjustments at the lower level. Our particular judgments about what to do can be criticized and modified, then, in the light of the more general moral principles we hold.

[4] See Parfit 1984; Daniels 1979.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Many, however, have complained that this exercise is extremely conservative.[5] While there may be some error in one's moral judgments that can be eliminated by pruning them in light of more general moral principles one also accepts, it is much more likely that the result of such an exercise will amount to little more than a "reshuffling of prejudices." Both our moral principles and our more particular moral judgments could simply "reflect class or cultural background, self-interest, or historical accident."[6] That a moral theory is not fraught with internal contradictions is hardly conclusive evidence for its acceptability.

In light of such complaints, it has been suggested that there are sources outside of moral theory which nevertheless constrain the content of that theory. And so we are to seek coherence among not only our particular moral judgments and our general moral principles, but also our background philosophical theories. Examples of such theories might include a theory of persons, a theory of the role of morality in society, a psychological theory of moral development, and general social theory.[7] It is hoped that these theories provide more resources for criticizing and reshaping the content of our given moral judgments than we had on the narrower conception.

---

[5] For example, see Hare 1975.

[6] Daniels 1979.

[7] Daniels 1979.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Let us consider a particular example in order to see how this criticism is supposed to work.[8] John Rawls famously accused utilitarianism of adopting a principle for distributing goods among several persons that is appropriate only for distributing goods among life-stages for one person.[9] And some Rawlsians[10] charge the utilitarian with ignoring an important feature of our philosophical theory of persons, for failing to shape one's moral principles in light of that theory.

Derek Parfit countered this charge by suggesting that the utilitarian holds a *different* philosophical theory of the person than does the deontologist, a theory that places less emphasis on interpersonal boundaries.[11] On this theory, the relations among various persons look more like the relations among various life-stages of the same person. The difference between the utilitarian and the deontologist, then, may stem from differences in their philosophical theories of the nature of persons.

If we feel strongly attracted to one of these competing theories of persons, reflecting on that theory will likely make one of the competing moral theories seem more plausible than it had before. And the same goes for our theories about moral psychology, moral development, social theory and the like. Not only can we adjust our particular moral judgments to fit the moral principles we find attractive, but we

[8] For a rich discussion of this debate, see Stern 1992.

[9] Rawls 1971, 26-27, 187.

[10] These are Rawlsians who have a somewhat different take on the matter than does Rawls himself—see Rawls 1975.

[11] Parfit 1984.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

can also modify both of these items in light of the background philosophical theories we accept. We have a larger body of thought against which to test our particular moral judgments. Charges of conservatism now look less plausible.

Now can the notion of wide reflective equilibrium help us address *our* problem of explaining immorality as a form of error? Can we show that immorality is a type of mistake by showing that the immoral hold incorrect background philosophical theories?

Suppose we are completely in the dark about what practical principles to accept. Suppose further that utilitarians and only utilitarians hold the reductionist theory of the person favored by Parfit. Now if we are firmly convinced both that the reductionist theory of the person is wrong, and that those who have the right practical views are very likely to have the right view about the nature of persons, then perhaps we can reasonably reject what the utilitarian says about what to do. And, of course, the same could be said about rejecting what the deontologist has to say if we happen to reject her understanding of the nature of persons. Thus it seems we can fairly explain the error in someone's practical view, if we find that he holds a deficient philosophical theory about some matter that would seem to bear weight on the determination of how to live. We can explain his error if we can show that the background philosophical theories he accepts are wrong, and that one is unlikely to have correct practical views so long as one's philosophical theories are off the mark.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

So can we show that the immoral are poorly positioned to grasp practical truths by showing that they hold incorrect background philosophical theories, perhaps in combination with holding poor information? Let me isolate two related reasons why I think we should be skeptical.

First of all, there is some doubt as to whether the philosophical theories in question are *independent* enough of the moral theory in order to serve as suitable *explanantia*. The philosophical background theories that shape and constrain our moral theories—the theory of the person, the theory of the role of morality in society, the theory of procedural justice, the ideal of a well-ordered society—these theories may already be so entrenched in questions of moral value, questions about what is important, that the link between philosophical theories and moral views is too tight to rule out heretical moral views fairly. We want to avoid moralism by showing rather than merely insisting that the immoral are mistaken, but the connections between philosophical theory and moral theory may be too close for us to escape moralism by relying heavily upon the former.

Furthermore, we may wonder whether it is even appropriate to use our philosophical views to explain why we have the moral views we do: doesn't the direction of influence run the other way? Nietzsche, for one, suspected that "the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

moral (or immoral) intentions in every philosophy constituted the real germ of life from which the whole plant had grown."[12] He speculated that

> if one would explain how the abstrusest metaphysical claims of a philosopher really came about, it is always well (and wise) to ask first: at what morality does all this (does *he*) aim?

The idea here is that one's philosophical theory—say, one's theory about the nature of persons—is driven substantially by one's moral views. We are reductionists (or non-reductionists) *because* we are consequentialists (or deontologists). The way in which the one can affect the other may be "unconscious" and "involuntary", but real all the same. And, of course, there is also a strand of Marxist thought that suspects that one's philosophical views are too much a product of class interest, of social factors that determine or constitute one's moral views, in order really to count as constraining our moral views. If anything, the direction runs the other way: it is our moral views that constrain which philosophical theories we adopt.[13]

If we nonetheless try to use the philosophical theories that seem plausible to us in order to narrow down what moral theory to accept, the worry is that we are simply using an obscure prejudice to justify an obvious one. If our philosophical theories are largely hazy and convoluted reflections of our moral views, then revising our moral views in light of our philosophical theories is like trying to get a better

---

[12] Nietzsche 1966 [1886], section 6.

[13] Martha Nussbaum forwards a similar conclusion in her debates with Bernard Williams about the concept of a person. See Nussbaum 1995, 86-131, especially p. 94.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

understanding of a visible object by examining closely the shadows it casts. It would be better to study the source of our philosophical theories, and the ways our moral views prejudice us toward some theories and not others, than to use those theories as a guide to explaining our moral views. Otherwise, we still seem to be reshuffling our prejudices.

But if we can somehow solve this problem, showing that our philosophical theories are independent enough from our moral views to avoid resurrecting moralism, a second problem looms large. We may wonder whether we can indeed explain the bulk of moral disagreement by looking towards differences in philosophical theory.[14] Perhaps some kinds of moral disagreement can be resolved through philosophical investigation. I even strongly suspect that many disagreements in moral *philosophy* could be resolved through convergence in other areas of philosophy.[15] It nevertheless seems unlikely that the people who we think hold incorrect moral views are *generally* the same as those who maintain philosophical theories we are prone to reject. It seems provincial and self-important for one who professionally philosophizes to suggest otherwise.

If we want to use the holding of certain philosophical theories as potential explanantia of moral error, then, we face a dilemma. On the one hand, the link between philosophical theory and moral view may be so close, that the only

---

[14] Rawls himself seems attracted to this criticism. See Rawls 1975.

[15] See G.E.M. Anscombe, "Modern Moral Philosophy" in Anscombe 1981.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

philosophical theories we will use are the ones we ourselves hold, hold largely because they fit well with the very moral views we are trying to explain. Thus the danger is that we fall back into the pitfalls of moralism. On the other hand, those philosophical theories that might escape the charge of moralism—this because they are more independent from the particular moral views in question—seem unlikely to explain very many of the differences among our views about how to live.

4.3 Looking Back at Advice

Sometimes we can explain someone's immorality by noting her lack of information; in other cases, we can explain it by showing that she has failed to bring her practical views into reflective equilibrium with sound philosophical theories. But frequently moral failure cannot be chalked up to such causes; too often people's practical views differ for reasons other than these traditional sources of error. Thus Williams' pessimism. If we hope to develop a theory of error, we will need to look elsewhere for additional resources to explain why people might fail to have the correct moral view. But where?

At this point, it is tempting to conclude that well-informed people can nevertheless coherently disagree about how to live, and that this shows that morality is not objective. Differences in practical view are no more than differences.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It would be a mistake, however, to throw our hands up so quickly. The case against moral objectivity looks overwhelming only if we adopt a monological view of deliberation. This is for two related reasons.

First, suppose we focus only upon the fact that people can disagree about how to live despite sharing the same information (and philosophical outlooks), or—as the case may be—despite the fact that they *would* continue to disagree even if they came to share the same information (and philosophical outlooks). The burden of proof now seems to be square upon the shoulders of the defender of morality to explain how, if there nevertheless are objective moral truths, such disagreement is possible. No other question seems to compete for our attention.

But, as a matter of fact, people not only *disagree* despite sharing the same information; they also come to *agree* when one person judges that another's advice is worthy of trust. And so there are multiple phenomena the metaethicist needs to account for. The opponent of moral objectivity cannot rest content waiting for the objectivist to succeed or fail at the task of explaining disagreement. Rather, the opponent of moral objectivity himself needs to argue that the kinds of indicator properties that trustworthy advisors possess do not really put one in touch with truths about how to live. The anti-objectivist needs to show that moral advisors are in fact *not* in a better position to figure out what to do than their advisees. Just as the anti-objectivist claims that the objectivist is rash to think that the wicked suffer from cognitive failures, so too can the objectivist claim that the anti-objectivist is rash to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

think that advisees are dupes. Each party to the metaethical debate needs to do some explaining, and it is unclear who has the harder task.

Second, not only is the burden of proof distributed more equitably—in that both objectivist and anti-objectivist now each have a task—but thinking about the qualities that make for a good advisor also brings into view resources helping the moral objectivist address *his* original task. For even if we cannot show that the immoral are ignorant or in a state of reflective disequilibrium, we may be able to show that the immoral lack other helpful qualities, specifically those qualities had by trustworthy advisors.

For notice that there is a difference between testing your moral views simply against your nonmoral views, and testing your moral views against the moral views of others. If I am wondering whether to ϕ, not only might I check to see whether the case for ϕ-ing essentially rests upon beliefs or philosophical theories I do not hold. I also might see what others think about my alternatives. As we saw in Chapter One, another person can advise me to act in ways I wouldn't otherwise have acted, even though our differences cannot be traced to differences in our beliefs about the nonmoral facts or in the philosophical theories we hold. And it can turn out that I trust another person's advice; I can decide to do what she says even though I would act differently if I were to approach the matter all by myself. The fact that *she* recommends ϕ-ing (or not ϕ-ing) plays a decisive role in determining what I do.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Now the features that we use to identify those whose advice is trustworthy may be qualities that both enable one to figure out what to do and yet be qualities that the immoral lack. The advisee thinks that his trustworthy advisor is in a better position to figure out what to do than he himself is. This advisee recognizes that he himself lacks qualities useful for figuring what to do, qualities had by the advisor. If having these qualities indeed puts one in a better position for figuring out what to do—as the advisee thinks—then it makes sense to wonder whether the immoral also lack these qualities. If advisors are trustworthy because they possess certain qualities, and the immoral generally lack such qualities, then it would seem that the immoral are not in a very good position to figure out how to live. It seems possible that the immoral occupy the same position as the advisee, the main difference being that the advisee recognizes his own limitations, whereas the immoral generally do not. So by recognizing that often an agent thinks about what to do by consulting the advice of others, a new range of resources for explaining moral disagreement as a form of error becomes available.

As we saw in Chapter One, the reasons an advisee might trust the advice of another are various. One of the more important reasons is that the advisee suspects that his advisor has undergone life-experiences that are likely to improve one's practical views. The advisee thinks that the advisor has led a sort of life which has strengthened her practical judgment. This may be because she has had done great things in her past; alternatively, she may have suffered evil at the hands of others, or

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

have faced hardships which the advisee has not. The advisee credits her recommendations with a soundness that his lack, largely because he recognizes that, relatively speaking, he is a bit naïve or provincial. He sees that she is in a better position than he is for determining what he should do.

Now notice how this differs from adjusting one's views about what to do in the light of someone else's views about the nonmoral facts. If I were seeking another's advice, I would not decide whether to trust her recommendations *directly* on the basis of whether she had good information, or the right philosophical theories. For I would not judge that her beliefs or theories were the correct ones, unless I myself held or adopted the same beliefs or theories. And, once I have adopted those beliefs or theories, there would be no grounds to think that her practical judgment is any sounder than mine: as far as I can tell, *both* our practical views have passed (or failed) the relevant tests.

But with life-experiences things are different. I can perfectly coherently think that undergoing some kind of life-experience improves one's practical judgment, even though I myself have not undergone it. I can count certain kinds of experiences as educative, even though I have not been so educated. This is in marked contrast to my thoughts about believing things: I will think that believing *p* will improve one's practical judgment only if I myself (come to) believe *p*. And so I can coherently and directly regulate my actions in light of the recommendations of those with different

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

experiences, even though I cannot coherently and directly regulate my actions in light of the recommendations of those with different beliefs or theories.

If our thoughts about how to live are enriched by educative experiences, then those who have not undergone such experiences will likely have inadequate views about what to do. Their inexperience, their naïveté, can make it more difficult for them to "see aright".[16] Lacking the sophistication and perspective one gains through educative experience, they often make poor decisions.

The fact that it can make sense to trust the advice of those with different educative experiences opens up a new range of resources for explaining immorality as a product of error. The immoral may be immoral precisely because they have not undergone the right range of experiences.

Perhaps people wrongly harm other people because they rarely face those affected in any significant way. Perhaps people lie because they have never experienced the gratitude one feels when one recognizes that someone else has disinterestedly put a stop to false damaging rumors. Perhaps people fail to prize generosity because things have always come too easy for them. Perhaps people are greedy for money because they have never experienced the threat of losing goods that money cannot attain. Perhaps people attack their neighbors because they have seen each other only in strained circumstances. Perhaps they pressure one another to

---

[16] Aristotle, *Nicomachean Ethics*, trans. Ross and Urmson, 1144a11-14.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

commit suicide because they have been isolated from much of what makes living worthwhile.

Defective moral views, then, might often stem from the lack of the right kinds of experiences, the undergoing of which improves one's thoughts about what to do and what not to do. The immoral may be in roughly the same position as the advisee, only that while the latter recognizes his limitations, the former does not.

Nonetheless, it is difficult to illustrate vividly the ways in which the lack of experience can cripple one's ability to grasp truths about how to live. It is difficult not because we are unfamiliar with examples of this kind of thing; rather, it is difficult because the examples we are familiar with tend to be personal. In part this is because you the reader probably do not know the people whose advice I myself am willing to trust, and I do not know the people whose advice you the reader are willing to trust. Good advisors often know their advisees well, and so it is difficult for me to point to an experienced person that makes a good advisor for all of us.

The other significant reason why it is difficult to highlight examples is that the sorts of experiences that improve the advisor's ability to determine what the advisee should do vary depending on the specific deliberative problem the advisee faces. The problems you face may be quite unlike the ones I do, and so you and I could benefit by heeding the advice of different people. No one person is likely to be in a position to tell each of us something useful.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But things aren't hopeless. Rather than pointing toward the ethical expertise of some moral guru as a convincing example, I will instead point out that each of us—if we are the least bit socially well-adjusted—has friends, family, counselors, clergymen, mentors, or coworkers whom we turn to from time to time. Those standing in certain social relations to us do so in part because they then can supply us with their thoughts about what they have learned from their experiences. A prime example of this is the way a person might advise her younger sibling about how to navigate adolescence, but the adult world affords still more difficult crises for which it is often very helpful to have someone who can share the fruits (or the bitter acid) of their joys and suffering.[17]

But perhaps a more respectable strategy of illustrating how the lack of life-experiences impairs practical judgment will draw comparisons between that thesis and other prominent work in moral philosophy on methods of moral inquiry. In particular, I propose that our life-experiences lead us to think that we should be just and beneficent and so on, just as other philosophers have proposed that reading literature can help us overcome our deficiencies. I develop this proposal in the following section.

[17] Cf. chapter 7 of Millgram 1997.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.4 Experience and Literature

Much of the excellent writing in recent moral philosophy has been devoted to thought about the relation between morality, philosophy, and literature; how literature fosters the reader's moral understanding in ways that scientific inquiry and philosophical argument cannot. There are parallels, I believe, between what some of the philosophers working in this field have to say about the educative power of literature, and what can be said about the educative power of the kinds of life-experiences we have been discussing.

I first want to look at a small part of a paper by Cora Diamond entitled "Anything but Argument?"[18] She maintains that argument is not the only way to make a convincing appeal for a moral view. Other ways of leading another to a change in moral view may include poetry, novels, satire, and other kinds of nonfiction *not* primarily composed of arguments. Examples are provided of each of these things to help remind the reader of these ways.

There are many aspects of this paper one could praise, but I will limit myself to a discussion of one. Diamond writes about how Dickens' particular way of describing how a child views the world—a description loaded with subtle moral aims—can alter the reader's conception about a child's life.[19] While before the reader

---

[18] Diamond 1991.

[19] Diamond 1991, 299.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

may have failed to notice the ways in which various events impress themselves upon children, or failed to think that any such differences between children and adults was very important, this now is brought to his attention, and the experience of reading Dickens sharpens his sensibility about these things. He probably now has a deeper sympathy, whereas before there was little.

Further, the change is due not solely to the reader's learning more facts about the child's situation:

> The moral significance of the kind of attention given by Dickens to the things he writes about is not a matter of its leading us to grasp facts of which we had previously been unaware . . . Dickens does not say: "Look at this: children do this and that, see thus and so, feel such-and-such, and these facts must be taken to be morally relevant." Rather his descriptions (not only *what* is described but the language in which it is) show an attention which *engages* us–if he is successful, and does not fail by getting the emotional tone off through sentimentality. Where he is successful, the description is not only enjoyable but can contribute to our lasting sense of human life, of what is ***interesting*** and ***important*** . . . Dickens's own view is that the investigation of facts, facts, facts *cannot* show us what we need in order to respond well to the world.[20] [***Emphasis*** mine]

We should want to say that this character of literature provides one with something over and above the facts of the situation described—even those facts involving thick ethical concepts—and it is this character which can be responsible for the change in moral attitude that can occur. Reference to this character or quality of the work seems necessary for any adequate explanation of what brings about the

[20] Diamond 1991, 299-301.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reader's moral responses to the work. The improvement in the reader's view cannot be fully characterized as the result of gaining new information.

Diamond's conclusion complements some of the findings of the present work. We saw that not all differences in practical view can be attributed to differences in information held. People who agree with one another about the facts can nonetheless have different attitudes about how to respond appropriately to such facts. The connection between belief and action is not straightforward.

But we also saw that not all such connections are to be regarded as equally sound. In fact, a person can think that another's view of what to do in light of the facts is sounder than his own untutored view. One reason why trusting another's advice may make sense is that her experiences are likely to have improved her judgment about these things. The significance of her experiences is not a matter simply of them leading her to grasp facts of which she had previously been unaware. Rather, her experiences have contributed to her sense of what is important, and this is one of the many things the advisee hopes to benefit from. Both literature and experience, then, can lead one to improve one's practical thoughts in ways other than by showing one the facts.

I also want to compare the present work to a paper by Martha Nussbaum about the relationship between philosophy and literature.[21] In this paper she argues

---

[21] "Introduction: Form and Content, Philosophy and Literature" in Nussbaum 1990, 3-53.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

that a proper investigation of the question about how to live will consider as many plausible answers as possible, and that thus we should study works of literature alongside works traditionally conceived as moral philosophy. Part of her argument consists in defending the idea that works of literature provide answers to this question that traditional philosophical writing cannot, and so that both styles make unique contributions to the investigation. Thus she is concerned to articulate how various elements of the form of, say, novels makes a difference in the ethical view they express, a difference inseparable from *that* form.

Much of her argument is given by example; it would be self-contradictory for her to try to say in a philosophy paper what she maintains a philosophy paper cannot say. Likewise, it would be foolish for me to try to summarize what she maintains a novel can tell us about how to live. To provide a flavor of what she has in mind, however, I should mention that Nussbaum thinks that the form of novels are specially and in some ways uniquely apt at conveying the incommensurability of value, the priority of the particular, the value of the emotions, and the relevance of uncontrolled happenings. Novels can convince us of these things in virtue of their structure, things that traditional philosophical writing cannot always defend as cogently. Focusing only upon traditional philosophical writing, then, would mean overlooking real options in our investigation of the question about how to live.

Similarly, we have found that one's experiences can assist one in thinking about how to live, assist one in ways beyond whatever help one receives in refining

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

one's actions in light of one's own beliefs. One gains wisdom and sophistication from living through difficulties; one's understanding of what is important deepens. This improvement in one's judgment could not always have be attained through further reflection. Reconciling the tensions and incoherences in one's thought is often very helpful, but sometimes one needs not a more orderly mind, but a broader one. New experiences can remedy a naïveté that is nevertheless immune to the kind of lessons one learns in the library, the seminar, and the colloquium. Both literature and life-experience, then, appear to be able to teach one things about how to live that traditional philosophical reflection cannot. Both literature and life-experience benefit one in ways other than showing one the facts, other than purging one's mind of incoherences.

Now I suspect that literature is able to enrich our moral judgment in this way *in part because* we are the kinds of creatures for whom our experiences nourish our practical thought. This claim is somewhat difficult to understand, much less defend. But I hope that, once pointed out, it has an intuitive plausibility. We learn how to live by reading about the experiences of characters largely because our own experiences can teach us about how to live. Literature has moral-educative powers largely because life-experiences have such powers.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.5 Autobiography as Advice

Now while philosophers have devoted attention to displaying the ethical education readers of novels and poetry receive, far less notice has been paid to the autobiography, but I think that this is unfortunate. Although nonfictional, the autobiography is in important ways unlike philosophy, science, even history; although narrative, it is unlike the novel, the drama. The autobiographer tells us about the experiences of his life, about how the events of his life changed him, and about what he now finds important or meaningful. In reading another's account of his or her own life, the reader can gain a kind of perspective and understanding on the question about how to live that is not duplicable elsewhere. In seeing how one person has chosen to live, one's own view of the possibilities, and of their various merits, can be widened.

The autobiography has both advantages over and disadvantages against the biography. The temptation to distort the truth, to hide one's warts, and to rationalize one's decisions plagues the autobiographer at every turn. Writing one's memoirs is often an exercise in self-delusion. This is one reason why semi-autobiographical fiction may prove to be the best vehicle for many individuals to relate the important contours of their own lives.[22]

But the author of the autobiography has a certain authority on his subject matter, an authority the biographer does not possess. While the biographer might be

---

[22] I thank Candace Vogler for pointing this out.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the better historian, the autobiographer is often the better source for understanding the educative results of the experiences of the subject's life. The reader of the autobiography can not only consider ways of life previously unimagined, but also peek into how ways of life she *was* considering might actually turn out. The reader can learn whether leading a certain kind of life is worthwhile, learn this from someone who should know.

Now it is true that *all* good writers bring what the experiences of their lives have taught them in touch with their writing. This is as true of novelists, essayists and poets as it is of autobiographers. But I want to suggest that each of these forms of literature can make its own unique contribution to the investigation of the question about how one should live, and that the autobiography has a special sort of role in this inquiry. The autobiographer can relate in an explicit way what her experiences have taught her, a lesson from which the reader can benefit.

Paul Fussell's writings, both autobiographical and critical, manifest and reflect these tendencies. As a young man, Fussell served as the leader of a rifle platoon in France during the Second World War. He saw most of his platoon die, he himself killed Nazi soldiers, and he was frequently living with the thought that this could be his last breath. He survived with a serious wound, and he went on to pursue a career as a professor of English and a literary and cultural critic.

In his writings, Fussell emphasizes both the limits and the virtues of autobiographical writings:

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> "War", says British author Nigel Nicolson, "is the activity of man about which more lies are told than about any other." This is why, like sex, it's not easy to learn much about it except by experience.[23]

> Vietnam veteran William Broyles Jr. articulates what all combat survivors know: "If you come back whole, you bring with you the knowledge that you have explored regions of your soul that in most men will always remain uncharted."[24]

> After scrutinizing closely the facts of the American Civil War, after seeing and listening to hundreds of the wounded, Walt Whitman declared: "The real war will never get into the books." Nor will the Second World War, and "books" include this one.[25]

The experiences the war veteran undergoes—even in a war as putatively just as the one Fussell fought—cannot always be communicated, not even through autobiography. The lessons learned often remain hidden; only simulacra of his fears and reliefs make their ways to the printed page.

Nevertheless, the autobiography remains perhaps one of the best vehicles for communicating the life-experiences of the combat survivor:

> Ever since my return to civilian life in 1946 I'd been recalling my experiences in war and considering their relation to everything else I knew. Did service as a young infantry officer in whatever time and place bring some special knowledge of humanity in relation to oneself? . . . Was my war unique or quite commonplace and hardly worth special notice? To see how widespread my experience had been, I sought out narratives by young literary-minded infantry officers with whom I could in some way identify myself. I didn't want fiction. I wanted testimony.[26]

---

[23] Fussell 1996, 160.

[24] Fussell 1996, 177.

[25] Fussell 1989, 270.

[26] Fussell 1996, 263.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> But the actualities of the war are more clearly knowable from some books than from others. The real war is unlikely to be found in novels, for example, for they must exhibit, if not plot, at least pace, and their characters tend to assume the cliché forms demanded by Hollywood, even the new Hollywood, and even if the novels are as honorable as Harry Brown's *A Walk in the Sun* and Mailer's *The Naked and the Dead*. . . . Despite undoubted success as engaging narrative, few novels of the war have succeeded in making a motive, almost a character, of a predominant wartime emotion, boredom, or persuading readers that the horrors have not been melodramatized. One turns, thus, from novels to "non-fiction", especially memoirs, and especially memoirs written by participants not conscious of serving any very elevated artistic ambition. The best are those devoid of significant dialogue, almost always a sign of *ex post facto* novelistic visitation. Because forbidden in all theaters of war lest their capture reveal secrets, clandestine diaries, seen and censored by no authority, offer one of the most promising accesses to actuality.[27]

We can see here the makings of an analogue to Nussbaum's argument about how the form of *novels* makes unique contributions to answering questions about how to live. The form of the autobiography also can convey aspects and features of our life-experiences that other forms of writing cannot.

Such considerations become markedly important when understanding and evaluating Holocaust narratives in particular. Then we might observe that

> Holocaust writers have been only too aware of the necessary difference between reality and imagination, and they have employed a variety of rhetorical devices to enforce the factuality or fictiveness of what I shall call the "contracts"—the patterns of literary expectation—that they establish with their readers.[28]

---

[27] Fussell 1989, 290-291.

[28] Foley 1982, 332.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

These differences, and the need to avoid diminishing what is all too real, imply that "some literary forms provide better frameworks that others for conveying the subject matter of the Holocaust."[29]

But even if the autobiography provides a special window into the lives of others, not all autobiographies are on an equal footing. Fussell takes to task those writers who never saw combat themselves, those military men who were at some remove from the front lines. He singles out, curiously, a philosophy professor at Colorado College, J. Glenn Gray, author of *The Warriors: Reflections of Men in Battle*. Fussell writes:

> But *The Warriors*, his meditation on the moral and psychological dimensions of modern soldering, betrays his remoteness from experience. Division headquarters is miles — *miles* — behind the places where soldiers suffer abject terror and madness and relieve the pressure by crazy brutality and sadism. Indeed, unless they actually encountered the enemy during the war, and encountered him face to face, most "soldiers" have little idea what war is like. Despite his sensitivity and intelligence, Gray's optimistic and congratulatory view of human nature never underwent testing on the line. [Underlines mine.] [30]

Gray knew quite a bit about life at headquarters; he knew little about seeing friends and enemies being ripped apart by mines and artillery. Being too naïve, the good professor lacked the sorts of experiences that might have improved his views about life.

---

[29] Foley 1982, 333. While Foley emphasizes the limitations of the conventional novel, she also argues that the traditional autobiography has its own problems, and that the diary may be more appropriate than either. For my purposes, I don't need to distinguish between the diary and the autobiography—both possess the same kind of authority over their subject matter.

[30] Fussell 1996, 292.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But it would be a mistake to think that each person's experiences are so unique and personal, and thus that there is no evaluative standpoint multiply occupiable. The form of Fussell's writings, for one, displays the ways in which human life is truly a common endeavor. Instead of describing the character of his own life in the trenches, Fussell instead often chooses to quote the memoirs or novels or poetry of those who lived in the trenches during the *First* World War: the similarities were sufficient to justify the substitution. Likewise, Fussell describes *The Great War and Modern Memory*, his historical-cum-literary opus about the British experience on the Western Front during the First World War, as "an act of implicit autobiography."[31] He could make use of such literary techniques only because of the ways in which life on the front affects men commonly.[32] Not only are there important differences among autobiographers—as the contrast of Gray with Fussell indicated—but there is considerable confirmatory overlap as well.

We can see in another's autobiography, then, the connections between the author's life-experiences and what he comes to think is important. If written well and honestly, the reader can learn something about what matters, learn it without needing to undergo such revelatory life-experiences himself.

---

[31] Fussell 1996, 266.

[32] See also Wiesel 1978, 200: "Have you read, reread, attentively read, the survivors' testimonies? They seem to have been written by one man, always the same, repeating a thousand times what you, the reader, even if you are his contemporary, will never understand."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

4.6 Reacting to our inexperience: some objections

I hope I have provided a somewhat clearer picture of why the lack of life-experience might tend to distort one's views about how to live. But it still would be a sham for me to insist that the bulk of immorality can be explained by noting the supposed inexperience of the immoral (in combination with noting whether they are poorly informed and in a state of wide reflective disequilibrium.) It would be a sham because we (I) don't have a firm grasp on how greatly and in what ways we benefit from the lessons life teaches us and is going to teach us. Would Agnes act morally if she had led a less sheltered life? Would she be less cynical if she had spent more time with the people she looks down upon? These are questions for which we can make at best only educated guesses. Each person's experiences are limited to the few strands of the vast fabric of human life with which she is familiar: who can say in what ways a person would change if she had undergone a wider set of experiences? Our inexperience even seems to far outstrip our nescience. There are so many facets of life we have not experienced, and so we are not in a position to tell whether the immoral would 'convert' if they were to lead richer lives.

We can react to our inexperience in two different ways:

> 1. We are not in a position to explain the bulk of immorality as an instance of error. Thus we have no reason to accept the conclusion that morality is objective.
>
> 2. We are not in a position to explain the bulk of immorality as an instance of error. But we see how such explanations might go. We cannot rule out

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> the possibility that these kinds of explanations work. Furthermore, we understand why we are not in a position to deliver such explanations, even if they do work. So we cannot reasonably deny (nor can we confidently affirm) that morality is objective.

Bernard Williams writes that he "cannot see any convincing theory of knowledge for the convergence of reflective ethical thought on ethical reality in even a distant analogy to the scientific case."[33] On his view, there is no explanation of most ethical divergence that shows it to be a product of error, and so he embraces the first option sketched.

But if it is reasonable to think that inexperience and naïveté can impair one's judgment about what to do, then it would be rash to embrace Williams' pessimism. Instead of thinking that there is no analogy to the scientific case of which Williams speaks, we have grounds to hope that immorality can be understood as a type of error, and thus that claims to moral objectivity can be sustained. While we may not be in a position to demonstrate whether morality is objective, we now can see better how such a demonstration would go. Reflecting upon the ways in which our life-experiences improve and distort our views about how to live shifts the burden of proof off the shoulders of the moral objectivist, making it more equally shared among all parties to the debate. Potential resources for explaining immorality as a kind of error seem huge: assessing whether they are sufficient, however, is exceedingly

[33] Williams 1985, 152.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

difficult. The second option, then, is more convincing than Williams makes it out to be.

A different sort of objection presses the claim that it would be plainly impossible for one person to attain the full range of experiences, and so my argument that immorality is a kind of error is incoherent. And indeed there is an element of unreality to the thought that one person could attain the degree of experience of which we are speaking. Each person has but one life to lead, and so each of us will necessarily miss out on much of the glory and grandeur of this world and its other inhabitants. But this abstraction is no more unreal than the notion of the absolute conception of the world usefully employed by Williams to defend the objectivity of science. The *Überwissenschaftleren* are no more flesh and blood that the Super-experienced. Both idealizations are useful devices which can potentially both aid and distort our thinking.

But a more serious problem for the proponent of moral objectivity stems from the fact that some types of experiences can be educative for some people, and yet distorting and corruptive for others. We saw this in Chapter One when discussing the effects of participating in a war—some veterans likely develop a more realistic view of the costs of war; others likely suffer a loss of moral sensitivity. It becomes difficult, then, to rely upon the property of having gone to war as indicative of knowledge about whether, in some new situation, one should war.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In order to get past this difficulty, we may need to look more closely at the details of the service records of the would-be advisors. Did they spend most of their time on the front line or at headquarters? Did they personally benefit from their service, receiving accolades and promotions, or were they instead consigned to the life of the typical soldier? Answering such questions may help us arrive at a better view about whether a person's experiences were likely to be truly educative, whether they indeed possess indicator properties for knowledge about what to do.

Alternatively, we may need to look more closely at the other effects of their experiences. Does it seem that their experiences have led them to do things which suggest they possess a coarse sensibility, steeled nerves that leaves them immune to the unfortunate plights of others? Or does it seem that they have a deeper sense of what is important, an understanding of pain and death that is neither shallow nor melodramatic and clichéd.

To figure out which veteran has the keener judgment—if any—seems to require the further exercise of one's own judgment. It takes no trivial degree of sophistication to determine whether another's judgment is sophisticated. And so the worry is whether relying upon the advice of an advisor can really improve one's own judgment: by the time one has figured out whom to trust, one has essentially already figured out what one will do. This suggests that relying upon the advice of an advisor cannot improve one's decisions over and above what one could have done exercising

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

one's own unaided judgment. And so the standpoint of the advisor might not provide an objective standard by which one can orient one's own activity.

No doubt the complexity of the decision about whom to trust often rivals that of the complexity of the decision about what to do. Sometimes one will have no more success in resolving how the multiple rival factors that bear upon the expertise of some potential advisor are to be balanced than one will have in resolving (by oneself) how the multiple rival considerations that bear upon what one should do are to be balanced. It seems we could use some advice about whose advice to take, some metaadvice. Unfortunately, I have little metaadvice to offer beyond the little I have already dispensed.

But things are not always so bleak. While we may not always be able to determine whose advice to take in situations where the same type of experience has influenced different people in different ways—not able to do this unless we are already sophisticated enough to answer the practical question for ourselves—many other kinds of situations may provide opportunities for benefiting from trusting another's advice.

Let me offer an analogy to suggest how this might work. It can be difficult for laymen to determine whether to trust the testimony of an expert witness when her testimony conflicts with the testimony of another expert witness. The layman can try to evaluate the credentials of each expert witness, to determine whether each sounds credible, and so on. But, ultimately, the layman may have no basis to decide which

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

expert witness—if either—is speaking the truth. It can seem that the layman would need to possess credentials roughly equivalent to those of the expert witnesses in order to determine which piece of testimony is the more sound, but by then the ex-layman would not need to trust what either says.

Nevertheless, the situation need not always be so futile. Even if a layman cannot evaluate conflicting expert testimony, perhaps a person with *some* credentials in the relevant field would be in a position to evaluate it. By knowing something about the pros and cons of the various research methods used, something about the kinds of training each of the testifiers has had, and so on, a person with some credentials may be able to draw some reasonable conclusions about which piece of testimony is the more sound. One needn't possess all the characteristics of an expert testifier in order to make such discriminations in a reasonable way. Rather, it seems that, even in difficult situations, one can figure out whose testimony is trustworthy if one's competence is just a step or two below that of the competence of those whom one is evaluating. (This obviously isn't a precise formulation of the relation between truster and testifier, but I hope that it is both comprehensible and plausible.)

Similarly, one may be unable to evaluate whether another's experiences have been educative if they are completely unlike anything one has experienced oneself. And so there will be times when it is difficult for one to reasonably determine whether an advisor's advice is worth trusting. But there likely will be other times when one *is* competent to evaluate whether someone else's experiences are likely to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

have been educative, namely, when there are moderate differences between what potential advisor and advisee have been through.

If this is right, then this opens up the possibility of metaadvice: Agnes may be unable to evaluate whether Carla is a good advisor, but Agnes may be able to evaluate whether Benjamin is a good advisor, and Benjamin may be able to evaluate Carla is a good advisor, so long as Agnes' experiences are somewhat similar to Benjamin's, whose experiences are in turn somewhat similar to Carla's. Thus Agnes may be able to reasonably trust Benjamin about whether to trust Carla's advice. These overlapping links could in theory compose a fairly long chain. Obviously, there are likely enormous difficulties and complications with putting this possibility in practice. But the possibility of such chains suggests that Agnes' inability to evaluate whether Carla's experiences are truly educative does not imply that there isn't a correct answer as to whether Carla is in a good position to figure out what Agnes is to do. (Unfortunately, the quality of such advice is likely to deteriorate as it becomes more and more difficult for the potential advisor (Carla) to evaluate the particulars of the potential advisee's (Agnes') situation.)

A related problem stems from the fact that there are some experiences that some people think of as *clearly* educative, and yet others take to be *clearly* corruptive. For instance, (it seems to me that) it would not be unreasonable for the virgin to trust the advice of someone with sexual experience about the merits of each way of living. Similarly, it *would* be unreasonable for us to trust the advice of the pedophile about

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the merits of his way of living, despite the fact that he has had experiences we have not. But what about the bisexual? the polygamist? Are *their* experiences educative or corruptive? Here we are likely to get a wide range of answers—some of which sound very confident—depending upon whom we ask.

This makes the role of experiences in explaining immorality more complicated than the role of information. Except for peculiar and paradoxical situations, more information rarely corrupts one's moral view; if it does distort one's moral view, it is generally because the information one gains is one-sided. Additional experience, on the other hand, might either improve or worsen one's practical judgment, and it is sometimes controversial which way things will go.

The explanations the moral objectivist can offer for why the immoral have the (putatively incorrect) views they do, then, may be ones *perfectly acceptable* to the immoral themselves. For instance, the sexually active moral objectivist may explain that the Shaker thinks that sex is wrong largely because she has never known for herself the joy of sex; the Shaker can agree with this explanation, but does not take it as showing that there is anything wrong with her view. She does not think that the explanation intended to debunk her view is indeed debunking. Similarly, the pedophile may explain that others think that pedophilia is wrong largely because they have never experienced what he advocates; we might agree that our views about the propriety of pedophilia very well might change as a result of engaging in such action—that is, we might think that engaging in pedophilia might distort one's

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

capacity for judging about the propriety of pedophilia—without thinking that his explanation debunks our view. The thing he points to as a putative indicator property of knowledge about such things is something we view as crippling.

Whether certain types of experience educate or corrupt, then, will not always be obvious. Explanations of why others fail to share our moral view, then, can be nearly as controversial as the moral view itself, *even if* this explanation is accepted by all as true. For the explanation of the other view was supposed to show it up as error, yet it is controversial whether it succeeds.

This naturally leads to the question: must the other person accept our explanation of his view—accept both its truth and the fact that its truth casts doubt upon his view—if the explanation is to avoid the moralism of which Williams speaks?

If we were doing *Ideologiekritik*, if we were seeking to articulate a Critical Theory whose primary goal is to lead a particular group of agents to cast off their self-destructive ideology, then it clearly would be necessary for the other to accept our contention about the educative nature of the experiences he lacks, or the corruptive nature of the experiences he has undergone. Anything less would constitute an inadequate critical theory.

Fortunately the standards for articulating a theory of error, the kind necessary to undergird claims to objectivity, are not as rigorous as those for Critical Theory.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

While the moral objectivist needs to genuinely show that the immoral are in some sort of error, he does not need to convince *them* of this. Here it is necessary to recall the conclusions of Chapter Three: a person may not recognize that she is in a poor position to grasp the reasons she has. A person can lack deliberative virtues even though she does not think of such qualities as virtues. And so a person can lack the kinds of experiences that would be educative even though she does not think of them as educative.

Of course, this in and of itself does not license us to charge those with whom we disagree as lacking appropriate educative experiences. The plausibility of such debunking explanations will need to be evaluated on a case-by-case basis. In the following chapter, part of my task is to take a particular case, and illustrate how to evaluate whether another person is in a good position to grasp what it is important to do.

4.7 Conclusion

In this chapter I have investigated whether the immoral are ignorant of truths about how to live. Traditional ways of demonstrating ignorance do not take the moral objectivist very far. Thinking about advice, however, draws our attention to deliberative resources we might otherwise easily overlook. Just as an advisee sometimes thinks that he himself is too inexperienced to know what to do, so too

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

might the immoral be too inexperienced to act well. It is difficult to say whether the lack of educative experiences is sufficient to explain the bulk of immorality, and so we are in a position neither to confirm nor deny whether the immoral are ignorant, and thus (from Chapter Two) whether proper moral views are explained by their content, a necessary condition for establishing the objectivity of morality.

Instead, the question about the objectivity of morality is beginning to look like an *empirical* rather than a purely philosophical question. Morality is objective only if those who have overcome all sorts of inexperience and naïveté, while avoiding all sorts of corruptive experience, would think that they should lead the moral life. And this is a question that the philosopher (qua philosopher) is not in a position to answer.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER FIVE

Do the immoral lack integrity?

5.1 NOT KNOWING, NOT CARING

So far, we have seen that the defender of moral objectivity must show that the immoral are not merely different from the upright, but in fact are dwelling in error. One way to do this is to explain why the *views* of the wicked are defective—or at least show that such explanations are possible. It seems that the defender of moral objectivity needs to argue that the immoral are *ignorant* about what to do, while the virtuous have *knowledge* about what to do.

When examining the available resources for arguing this claim, we found, not surprisingly, that it is difficult to say whether the immoral would in fact change their views about what to do were they to undergo the full range of educative experiences. This is in large part due to the fact that no single individual *can* undergo all the educative experiences that might improve one's moral views. There is also the problem of identifying which experiences are indeed educative, and which are corruptive. Thus, while we have a schema by which we might be able to explain why the immoral are ignorant, we cannot say for sure that such explanations in fact exist. Moral objectivity is a *coherent* hope, but still only a hope.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I now want to consider another way we might describe the kind of error or mistake stemming from the lack of experience. Up until now, we have been speaking about whether we have grounds for saying that the immoral fail to grasp moral truths. The thought we would like to establish is that the upright know what they ought to do, while the immoral do not. But perhaps for a large number of cases, it would be more accurate to say, not that the immoral don't know what they ought to do, but that they *know all too well* what they should do. It seems that many wrongdoers know that they shouldn't do what they nevertheless do. These folks do not suffer from simple ignorance, but from a lack of integrity. Their actions are out of line with what they know; their selves lack unity, wholeheartedness, integrity.

The possibility of showing that the immoral lack integrity opens up a broader and more promising strategy for the defender of moral objectivity. Attacking from one side, he can still argue that many of the immoral lack knowledge about what they ought to do, this because of their ignorance of nonmoral facts, lack of imagination, incoherence between their moral and philosophical theories, inexperience, and so on. Attacking from the other side, he also can argue that the rest of the immoral know what they ought to do, but lack integrity—they fail to live morally because they are self-deceived, akratic, or otherwise practically irrational. The defender of moral objectivity, then, can try to put his opponent in a vise, pressing in from one side with the argument that many of the immoral lack moral knowledge, while pressing in from the other with the argument that the rest lack integrity. His goal is to leave no room in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the middle, no room for the immoral person who is both wholehearted and free of ignorance.

We will have to take several steps in order to see whether this strategy is successful. First of all, what is integrity, and why is it important? Second, how is lack of integrity a form of error? That is, even if we could show that (many of) the immoral lack integrity, how would this aid the defender of moral objectivity? Finally, and most difficult, do the immoral indeed lack integrity? What grounds do we have for thinking that they do? How strong a case can the moral objectivist make?

5.2 Collecting Your Thoughts

I have suggested that we often do wrong, not because we don't know that it is wrong, but we do it despite having this knowledge, and that this displays a lack of integrity. But what is integrity?

We could try to answer this question simply by enumerating the various forms of integrity, and the ways in which a person can fail to exhibit integrity. As I mentioned above, both the self-deceiver and the akrates display a lack of integrity. This suggests some sort of link between integrity and rationality. But not all forms of irrationality are expressions of a lack of integrity; for example, a person who obstinately retains her first formulation of a theory despite new strong conflicting evidence may be irrational, but she does not thereby exhibit a lack of integrity. The

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

forms of irrationality that David Pears has called perversions of reason—those habits of human reason itself that often lead reasoners to falsehoods—are not examples of the sort of thing that betray a lack of integrity.[1] Integrity is not exactly the same as competence at processing information.

If we think of the person of integrity as principally one who is neither self-deceived nor akratic, then this is because these latter forms of irrationality suggest, at least metaphorically, some mental division or fragmentation in the person who suffers from them. This mental division is opposed to the kind of psychic structure possessed by the person of integrity, as a surface investigation of the etymology of the word *integrity* confirms. One who has integrity is wholehearted, is honest with himself, does what he in fact thinks is best (whether or not it is the same as what is moral), and—though how further to understand this is debatable—is of stable character. The person who lacks integrity is often halfhearted, deludes himself, does not do what he thinks is best, and is of wavering character.

Integrity, then, is a kind of wholeness of personality. The psychological states of the person of integrity have their full rational effect upon other elements of his psyche. He does not deceive himself, for he does not avoid acknowledging the implications of available evidence as to how things are. He does not act against his better judgment, for he gives all considerations (what he takes to be) their due weight

[1] See Pears 1984, 9. See also Nisbet and Ross 1980.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

when forming that judgment, and he then sticks by it.[2] Although rich and multi-faceted, the various aspects of his personality complement rather than conflict with one another. The person of integrity is someone whose "personal expression" conforms to his "psychological reality."[3] The person of integrity is one with himself.

5.3 Looking for the Lukewarm

As so far described, the person of integrity sounds like a pretty good guy. In fact, the opponent of moral objectivity will likely suspect that he is too good. The worry is that, even if it is shown that the immoral lack integrity, this will not show that the immoral are making any kind of error, at least not any kind that can be understood from an *external* perspective as problematic. It might seem that to show that the immoral lack integrity is no more damaging than to show that they lack the virtue of temperance, or perhaps even justice. Which is no more than to say that the immoral are immoral, hardly a conclusion with stunning implications. If integrity is already a moral quality, then nothing follows if we were to show that the immoral lack it.

---

[2] See Williams 1981, 40-53 on the commitments that the person of integrity sticks to.

[3] Robert Grudin (1990, 73-74) defines integrity as "1. an inner psychological wholeness; 2. a conformity of personal expression with psychological reality—of act with desire, of word with thought, of face with mind, of the outer with the inner self; and 3. an extension of wholeness and conformity with time, through thick and thin. Though integrity can be, and must be, expressed in individual actions, it is not fully realized except in terms of continuity."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

But I want to make plausible the thought that integrity is not itself a moral virtue, that it is independent enough from questions of morality such that it is an interesting question whether the immoral lack integrity. Integrity is a quality that is instead rather formal, one that concerns the relations among a person's mind and actions.

To illustrate this formal character, consider Aristotle's sixfold division of states of character: divine excellence, human excellence, continence, incontinence (*akrasia*), vice or self-indulgence, and brutishness.[4] Aristotle says little about the first and last of these—these states are not ones humans can really possess—but he describes the other four in some detail. The virtuous person, the one who is not even tempted by what is bad, surely has integrity. But so does the vicious person, for "the self-indulgent man, as was said, has no regrets; for he stands by his choice . . ."[5] That is, there is no internal conflict within the soul of the self-indulgent man; he exclusively pursues wine, women and song with no twinge of shame. Although morally blameworthy, it seems that we cannot launch an internal critique of his ways. His actions are in complete accord with all of his beliefs and desires about how to live; "vice is in accordance with choice."[6] So both the virtuous and the vicious are paragons of integrity; integrity is not (only) a morally positive quality.

---

[4] Aristotle, *Nichomachean Ethics*, trans. W.D. Ross and J. Urmson, VII.1.

[5] Aristotle, *Nichomachean Ethics*, trans. W.D. Ross and J. Urmson, VII.8.

[6] Aristotle, *Nichomachean Ethics*, trans. W.D. Ross and J. Urmson, VII.8.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The incontinent person, by contrast, is the one who most clearly displays a lack of integrity. He thinks that it is best not to ϕ, and yet he ϕs anyway. His judgment that it would be best from him not to ϕ is not efficacious, shut out from that part of him responsible for acting. His self seems to be fragmented, his judgment and action out of step with one another.

But what about the continent person, the one who, while tempted to do other than what he judges best, nonetheless acts is in accordance with his judgment? Does he lack integrity because his appetites attempt to pull him against his judgment? One might think that

> [a] man with incompatible desires . . . is less a single agent. Rather, he is two (or more) agents competing against each other.[7]

Or does the continent person instead display integrity, since, after all, his action issues from his best judgment?

It would be too restrictive to demand that a person not have appetites that conflict with his final judgment (or indeed one another) if he is to display integrity. The person with integrity is not fragmented, but he is not monolithic either. The person with integrity will neither ignore the appetites that try to pull him away from his judgment—he consults them *in* forming his judgment—nor count them twice. He assesses their worth in light of all he knows and understands, allowing them to

[7] Neely 1974, 39 as quoted by Darwall 1983, 103.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conquer or be conquered by the other considerations operative. The continent person, it seems then, does not fail to display integrity.

Acting against one's best judgment is of course just one way one can fail to display integrity. Willfully ignoring available evidence in making that judgment is another, and there are many other shapes such failure can take. But these kind of failures are not identical with moral failures; whether or not it is plausible, one could *coherently* argue that the person of integrity would reject morality. Indeed one suspects that those who reject morality think of themselves as being less self-deceived than those who embrace it. Integrity is instead something of a formal quality, recognized by both friends and enemies of morality alike as something worthwhile.[8]

One writer claims, perhaps only a bit too strongly, that from the fact that someone is a person of integrity "nothing whatever can be *deduced* about any *particular* kind of behavior on the part of that person" (*emphasis* mine).[9] Perhaps nothing particular can be deduced about the behavior of the person of integrity, but we will expect that her future behavior will be consistent with her past, and that it will be informed by her best understanding of both herself and the world. This is surely not logically inconsistent with a life of immorality, and so the opponent of moral

---

[8] See Rawls 1971, 519-520, who calls integrity a "virtue of form."

[9] Taylor 1981.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

objectivity should not balk at this investigation on those grounds. Thus if the immoral lack integrity *as a matter of fact*, that would be both interesting and relevant.

But would it help the moral objectivist? Even if the immoral suffer from a lack of integrity, and even if this is something they themselves would find troubling, how does this show that the immoral are thus in error, the kind of error that would vindicate claims of moral objectivity?

Consider again the strategy of Chapter Four, the strategy where we try to show that the immoral are likely to be *ignorant* of truths about what to do. Success in this project would show that the immoral are cognitively cut off from the truth, that their views about what to do are incorrect. The immoral, then, would not be merely different from the moral; they would be mistaken.

Following this new strategy, we are not concerned with the case where the immoral person is cognitively cut off from moral truth. Rather, we are hypothesizing that these moral truths, in some sense known by the immoral person, nevertheless have failed to resonate throughout his personality. *He* is not cut off from moral truth, but certain parts of his personality—those that are aware of moral truth—are cut off from other parts of himself. Even if his immorality cannot be explained by showing that he is in a poor position to acquire moral truth, perhaps his immorality can be explained by showing that he lacks integrity, that his psyche is fragmented, such that he *does* know moral truth, but it is not allowed to have due influence upon what he

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

does. This would strongly suggest that there are indeed moral truths, but that the immoral person refuses to live his life in light of them.

Any plausible and realistic strategy, however, cannot merely suppose that those who are immoral nonetheless *must* know that their actions are not appropriate and that *therefore* the immoral suffer from a lack of integrity. This would be getting things back to front, assuming what we are trying to show, guilty of the kind of moralism Bernard Williams so rightly identifies. Instead we would need first to show that the immoral suffer from a lack of integrity, that their fragmented lives are an unripe environment for the presence of moral views to translate into moral action. It will be difficult to establish definitively that the immoral indeed have moral knowledge, views which are cut off from the rest of their personality. Rarely will they explicitly verbalize the fact that they have such views, or that their lives lack integrity. Often we will instead have to look for signs of the presence of these views, clues that their immoral actions are not fully indicative of all that they think and feel.

5.4 Fragmented Nazis

In order to show why this might be an accurate psychological profile of the immoral, I will discuss some curious features of the moralist's favorite example of evil, the Nazis. I do this not to browbeat or bully anyone into moral objectivism. Rather, I think it will be genuinely helpful. Recall Socrates' funny little argument in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the *Republic*, where he says that if we want to better understand what it is for a man to be just, we should first examine what it is for a city to be just, for the justice in a city will be larger and therefore easier to spot. Even if this is a bad argument, I think a somewhat similar strategy may nevertheless be appropriate for us. Qualities that tend to characterize most immoral people may be exaggerated or magnified in those as wicked as the Nazis. By understanding as best we can the nature of Nazi evil—difficult as it may be—we will be in a better position to grasp some more general truths about the causes of immorality.

In thinking about Nazi evil, we also need to be aware that their wickedness is likely to be explained by multiple factors. For instance, while it is surely true that the Nazis would not have done as much evil if they had not held incorrect views about biology, race, heredity, and the like, other factors also contribute to the explanation of their immorality. The totality of their evil is unlikely to be completely explained by noting their bizarre science. Understanding how other considerations also contributed to Nazi immorality may help us see that the same considerations contribute to more everyday forms of immorality.

Furthermore, it is likely that any single explanation will work for understanding the wickedness of at best only some perpetrators. Holocaust studies are reflourishing, due in part to the controversy surrounding the explanations put

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

forward by one author.[10] We can disagree about the scope of any such explanation, but it is likely that each of the major competitors will have some place in the final analysis. I focus upon those explanations that seem to me to be the most convincing, those which purport to reveal which factors were necessarily involved in enabling genocide. But, even as a philosophical reflection upon the historical and social psychological work of others, what I have to say here is radically incomplete.

Nevertheless, let us look at a few strands of explanation. Psychiatrist and social psychologist Robert Jay Lifton forwards the claim that the key to understanding how fairly ordinary German doctors came to commit genocide is the psychological principle of "doubling".[11] According to Lifton, the self of the Nazi doctor split into two subselves, such that psychological material conducive to killing motivated most of his life, while the part of his self associated with healer and father was rendered largely ineffective, though not completely jettisoned. That is, the Nazi doctor still thought of himself as a healer, as a physician, as a loving father, and even vigorously pursued projects to reinforce these aspects of himself, but this psychological material rarely interfered with the operations of his Nazi self, even when such interference would have made sense, as it often would have.

Doubling is to be distinguished from the psychoanalytic concept of splitting or dissociation, which tends to mean that the split-off portion of the self ceases to

---

[10] Goldhagen 1996.

[11] Lifton 1986.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

respond to the environment, as well as from schizophrenia and multiple personality syndrome, radical divisions that tend to afflict individuals throughout their lives.[12] Rather, doubling is a more temporary phenomenon, an adaptation to an extreme *social* environment, one in which an individual's anxieties and weaknesses as well as certain institutional pressures bring about a "dissolving of psychic glue."[13]

Another line of explanation complements the previous one. Literary critic and scholar Tzvetan Todorov meditates upon the fact that both survivors of Auschwitz and contemporary historians repeatedly draw our attention to the fact that all of the camp guards demonstrated remarkable behavioral inconsistency.[14] Helpful at one moment, vicious at the next, the lives of those responsible for this great evil were, as Todorov puts it, "fragmented". Elements of goodness, of caring, of humanity would leak out from time to time, whether it be something as dramatic as saving a prisoner marked for the gas chamber, or as mundane as taking special care for an inmate's momentary need. One camp survivor remarked that the most horrifying thing about the Nazis "is the good inside them; the saddest thing is that they have so many good qualities, so many virtues."[15] The worst among us nevertheless seem to retain some virtuous aspects; conversely, our good qualities do not ensure that we will completely

---

[12] Lifton 1986, 420-423.

[13] Lifton 1986, 422-423.

[14] Todorov 1996.

[15] As reported by Todorov 1996, 142-3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

avoid wrongdoing. Much more often, however, these good parts of the camp guard's psyche were split off from the subself generally operative. Their more ordinary thoughts, feelings, wants, and plans did not prevent them from usually behaving like monsters.

I shall discuss some ways in which this doubling or fragmentation manifest itself in the lives of Nazi wrongdoers. We will see that they suffered from self-deception as well as other conditions signaling a lack of integrity. I shall also consider some explanations for why this fragmentation occurred. We will see that the Nazis could remain psychically divided in part by avoiding face-to-face experiences with those whom they wronged. Finally, I shall briefly argue that the sources of Nazi evil are not completely anomalous, that we can better understand the origins of everyday immorality by understand the origins of Nazi immorality, and that many (though probably not all) forms of immorality stem from a lack of integrity.

5.5 Repressing humanity (your own and others)

Raul Hilberg argues that the German bureaucrat coped with his moral inhibitions through "a mechanism of repressions and a system of rationalizations."[16] These repressions and rationalizations manifest themselves in ways both subtle and severe.

[16] Hilberg 1985, 278.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

First, we need to understand the nature of the psychic division of the typical Nazi evildoer. Those who developed the capacity for evildoing tended to change in ways that made this transformation easier. Before being directly involved in murder, they were surrounded by an environment where stories about such killing were in circulation, but were neither clearly documented nor avowed.[17] Perhaps because of the shocking nature of the very subject matter of these accounts, many were unsure whether to believe that these stories were accurate.[18] People tended to both believe and not believe the rumors of wrongdoing. (This is true not only of many potential killers, but also of many potential victims.) Lifton describes this as a kind of "middle knowledge", a situation in which people both know and do not know about the killing, or they know but do not act as though they know.[19] Middle knowledge about what is going on influences how one lives, yet it does so in ways that makes it easier for one to avoid confronting the more difficult aspects of what one is up to. Having middle knowledge of these atrocities makes it easier, in some ways, to allow oneself to become more involved in them. One Jewish prisoner doctor explains the experience of the Nazi doctors in this way:

> The fact is that if you do something that is totally unbelievable, and you are incapable of believing, you don't believe it. . . . The gas chambers, . . . the

---

[17] Hilberg 1985, 278-279.

[18] Hilberg 1985, 244.

[19] Lifton 1986, 489.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> houses that the crematoria had, . . . brick houses, windows, curtains, white picket fences. . . . Nobody would have believed that.[20]

There are many matters about which we are unsure, about which we do not have any firm beliefs, but these are generally not cases of middle knowledge. One can perfectly lucidly review evidence and determine that it does not warrant any conclusions. In such cases, it seems appropriate to say that one does not have any belief about the matter in question.

Whereas in cases of middle knowledge, it seems more appropriate to say that the inquirer *both* believes that the stories of killing are true, and believes that the stories of killing are not (could not be) true. In these cases, a Freudian understanding of irrationality seems to be largely correct. The inquirer believes that the stories he hears are not true, because, despite the modest reliability of his sources, the inquirer at some deep level wishes that they were not true. The wish that things were otherwise fosters the belief that things *are* otherwise, and so the inquirer's mind divides as his thoughts become less rationally responsive to one another.

The division in the mind of the potential killer makes it possible for such a person to eventually become accustomed to the killing that is going on. One blocks out the parts of one's self that would naturally oppose the social pressures that encourage complicity. Gradually, one feels more and more at home in an increasingly

---

[20] As reported by Lifton 1986, 447.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

murderous atmosphere, and the atrocities that surround one eventually seem like, as one witness of Auschwitz put it, part of the weather.

The Nazis' inconsistent beliefs often lead to inconsistent actions. As noted above, Nazi doctors (and other camp workers) treated the prisoner-patients in radically incongruous ways. One Nazi doctor described Auschwitz as a "schizophrenic situation". SS doctors sometimes would give pregnant women a double food ration, then send them to the gas chambers the following day. Prisoner doctors say that the Nazi doctors were "different persons" when doing selections. The fragmentation was most evidenced by the fact that doctors tended to rush away quickly after doing the selections, as well as turning operations over to underlings whenever possible—they wanted to distance themselves from what they were doing.[21]

One Auschwitz survivor writes of Josef Mengele:

> He was capable of being so kind to the children, to have them become fond of him, to bring them sugar, to think of small details in their daily lives, and to do things we would genuinely admire. . . . And then, next to that. . . . the crematoria smoke, and these children, tomorrow or in a half-hour, he is going to send them there. Well that is where the anomaly lay.[22]

To at least some survivors, it did not seem that Mengele was pretending to play with the children merely in order to gain their confidence. Prisoner doctors who worked with him spoke of Mengele as truly having a "split personality":

---

[21] Lifton 1986, 210-11.

[22] Lifton 1986, 337.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> The double, that is to say he [Mengele] had all the sentimental motions, all the human feelings, pity, and so on. But there was in his psyche a hermetically closed cell, impenetrable, indestructible cell, which is obedient to the received order. He can throw himself in the water to go and save a Gypsy, try to give him medication, . . . and then as soon as they are out of the water, . . . tell him to get in the truck and quickly off to the gas chamber.[23]

The fracture in Mengele's self was probably more severe than that of the typical evildoer, but this kind of behavioral inconsistency nonetheless seemed to be the norm among Nazi physicians.[24]

The doubling process further manifests itself in the way Nazi doctors isolated their work from their family.[25] Wives and children rarely visited Nazi physicians at the camps. Many of the camp workers expressed their desire not to taint their families with their Auschwitz work: the family represented ordinary values, purity, compassion, a safe haven in which the non-Auschwitz self could reaffirm itself. Spending time at home, away from the camps, enabled them to continue to think of themselves as decent human beings. It kept them in touch with the values associated with normal life. Generally speaking, keeping this aspect of themselves alive but

---

[23] Lifton 1986, 374-375.

[24] Cf. Hilberg 1985, 289: "When the trials of war criminals started, there was hardly a defendant who could not produce evidence that he had helped some half-Jewish physics professor, or that he had used his influence to permit a Jewish symphony conductor to conduct a little while longer, or that he had intervened on behalf of some couple in mixed marriage in connection with an apartment. While these courtesies were petty in comparison with the destructive conceptions that these men were implementing concurrently, the 'good deeds' performed an important psychological function. They separated 'duty' from personal feelings."

[25] Lifton 1986, 211. Hilberg 1985, 241.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

separate from their work life was crucial if they were to continue to function in the camps.[26]

Camp life displayed it own puzzles, however. The most prominent example of doubling in Nazi physicians was what Lifton calls "the healing-killing paradox". Nazi doctors played a crucial role in the Third Reich's killing machine, beginning with the forced sterilization programs, continuing with the euthanasia of the mentally and physically ill, leading up to the T4 killing project and the institution of mass killing in prisoner camps. The physicians had to see their killing as necessary for the healing of the Aryan race, in order for them to feel that what they were doing made any sense. To the extent that they entirely accepted the Nazi biomedical vision, the paradoxical nature of this relation between healing and killing vanished, resembling instead the killing a virus in order to heal the body.

Yet historians and social psychologists see in the lives of most Nazi doctors peculiar features that we do not find in the lives of most practicing immunologists, features or clues that suggest that while on the one hand they intended to carry forth the Nazi program, they also had buried within them reservations about whether this is what they really were to do.[27] Individual doctors and the Nazi organization had to

---

[26] Hilberg 1985, 281-282.

[27] Laurence Thomas—whose work on Lifton's account of evil I discovered long after I wrote the bulk of this chapter—maintains that "Lifton's account of doubling would not seem to be at all explanatory. That is, reflection upon the features that he takes to be generally characteristic of the phenomenon gives us little or no insight as to how the phenomenon could occur." As I go on to show, Lifton in fact does provide with material for understanding how doubling occurs. See Thomas 1993, 97.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

struggle to maintain the fraud that genocide was a medical solution to a medical problem.

To facilitate this fiction, SS physicians were addressed by "Doktor" instead of by military rank, as was typical, in order to reinforce the connection between healing and their work in the camps.[28] Gas, gassing personnel, and victims were transported to gas chambers in trucks marked with red crosses.[29] These and other similar routines bathed the ongoing killing in an environment typically associated with healing.

Selections in particular, the most common and for many the most difficult work a typical camp doctor performed, fully embodied the paradoxical nature of his role as well as the ways in which the organization made it possible for him to adopt it. In a period of two or three seconds, physicians "selected" for the gas chambers those too "ill" to live and work—hardly enough time for a professional doctor to exercise his diagnostic skills. It was clear to any observer that it was completely unnecessary for highly trained physicians to conduct the selections—women, children, the old, the emaciated were generally sent to the smokestacks; injury-free men with muscle and color sometimes were not.[30] Deviations from this pattern were either politically motivated—orders from above would determine how many lived—or completely random; they were not medically based. Yet the whole operation could run without

---

[28] Lifton 1986, 452.

[29] Lifton 1986, 431., Hilberg 1985, 248.

[30] Hilberg 1985, 245-246.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

German resistance only as long as the killing remained associated with the healing of the *Volk*, and the medicalization of the mass murder preserved this association.

Several features of the selections process made it psychologically possible for a fairly ordinary German physician to perform them. It should not be forgotten that the typical German was anxious about his own death, both because of the war and because insubordination was not treated lightly. Wielding power over the lives of others is one way individuals cope with their own feelings of vulnerability. Second, the new doctor often was welcomed to the camp and initially monitored by the more famous physicians, and the camaraderie that developed between him and those whom he professionally respected made it psychologically possible for him to accept the ways of the latter. The social pressures on the new camp doctor were significant, and, not wanting to be an outcast, the new doctor came to think that performing selections was what doctors did (which is not to deny that he *also* knew better).[31]

Furthermore, it is important not to overlook why the selections process was highly routinized. Even though a doctor typically spent only a few seconds examining a prisoner before making a decision, the rest of the process was heavily structured so as to lend it a certain aura of respectability. Selections might happen at the same time of day every day; prisoners were made to line up and were repeatedly counted, life in the camps became structured around the idea of and the "need" for

[31] See Milgram 1974.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

selections. Although his first experience of selections was shocking and disturbing, the typical physician quickly got used to the whole thing:

> When you see a selection for the first time—I'm not talking only about myself, I'm talking about even the most hardened SS people, . . . you see . . . how children and women are selected. Then you are so shocked . . . that it just cannot be described. And after a few weeks one can be accustomed to it. And that change cannot be explained to anyone. But it is the same phenomenon that takes place right now in terrorists, in relation to close terrorist groups.[32]

No doubt the ritualization and the institutionalization of the selection process helped to make it psychologically possible for physicians to kill millions of people. It contributed to their feeling that they were just tiny cogs in a huge machine, not really responsible for the consequences of their actions. A lack of routinization would have made it harder for them to rationalize what they did.[33]

The widespread use of alcohol also made it possible for German doctors to kill despite their self-identification as healers. Inebriation facilitated the psychic numbing necessary for Nazi physicians to do their job. Doctors would drink together in the evenings, share with one another their reservations about their activity, even denounce the whole killing project. This was probably the only setting where such reservations could be voiced.[34] But these drunken antics didn't interfere with what the doctors *did*; rather, it seemed that the opportunity to vent their reservations made it

[32] A Nazi physician who worked at Auschwitz, as reported by Lifton 1986, 197.

[33] Hilberg 1985, 126, 243-249, 288.

[34] Hilberg 1985, 279-281.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

easier to kill. New doctors would raise questions about what they were doing; the older doctors would then rationalize what was going on. Lifton suspects that the opportunity to raise and combat doubts helped to suppress aspects of the self that opposed the killing, nurturing the development of a new Auschwitz self.[35] The Auschwitz self could think of selections work as legitimate work for a physician. Pretty soon, the new doctors became as resigned to their actions as the older ones.

Thus they came to understand the killing as a form of healing. For instance, when an entire block of prisoners suffering from extensive diarrhea or typhus was sent to the gas chamber—not an uncommon event—Nazi doctors could think of this action as promoting the health of the rest of the camp. Doctors could give children overdoses of sedatives, convincing themselves that otherwise the children would be too restless for their own good.[36]

Sometimes, the self-deception about what one was doing was more severe. One prisoner doctor told Lifton:

> I couldn't ask [Dr. Fritz] Klein, "Don't send this man to the gas chamber", because I didn't know that he went to the gas chamber. You see, that was a secret. Everybody knows the secret, but it was a secret. If I said to him, "Herr Doktor Klein, why should you send this man to the gas chambers?", I suppose that he would say, "Gas chamber? What do you mean?"[37]

---

[35] Lifton 1986, 195-6.

[36] Lifton 1986, 54.

[37] Lifton 1986, 202-3.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This kind of pretending not to know about what one was in fact doing contributed to "an atmosphere of denial" and ambiguity in which there were fewer pressures to confront fully the consequences of one's actions.[38]

Individual doctors also often pursued projects that made it easier for them to see themselves as physicians. Frequently they took advantage of the availability of human subjects in order to perform medical experiments, this in part to bolster their self-image as physicians forwarding medical science, even though the results for the victims of these procedures was disastrous. Despite the terror these physicians inflicted on their experimentees, the opportunities the camps provided for faux research actually made it psychologically easier for many of the Nazi doctors to continue to see themselves as doctors. The victims of these experiments were thought to be nearly dead anyway; the doctors could convince themselves that the research results would benefit future (Aryan) humanity; unlike in their work doing selections, they could play with the tools of their trade (e.g. scalpels, thermometers), they could write up their findings. The camp experiments, although in many ways the most gruesome feature of the Nazi experience, made it possible for physicians to contribute to mass killings. Opportunities for research helped ease the psychic tension in being both healers and killers.

[38] Cf. Hilberg 1985, 281-282.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

For example, one Nazi physician, Dr. Horst D., expressed reservations to his superiors about participating in the killing. When a leading German neuropathologist appeared at the camp to conduct research, Dr. D arranged to spend some time studying correlations between the medical histories of patients and features of their brains in order to learn more about specific diseases. Dr. D. could now see the killing machine as an opportunity for medical research; he could suppress his doubts by thinking of the gas chambers as a necessary step for his medical work. Kidding himself about the supreme value of science, Dr. D was able to become a murderer.[39]

The Nazi doctors had to work hard in order to maintain the fiction that their work was medical in nature. At some level, many of them came to believe that it indeed was. But at the same time, they also knew that it was not, and we can confidently ascribe to them this inconsistency only after we have understood the presence of, and also the need for, these factors which served to distract their attention from the nature and consequences of their work, clues which suggest that they lacked integrity. Though typically very anti-Semitic, Nazi doctors did not fully and lucidly swallow all of the implications of the equation of healing with killing. Drinking, fear, the demand for routinization, the need for "research", the role of professional camaraderie—these and other features of their lives suggest that the Nazi doctors wanted to avoid thinking about what they were doing. These features are some of the

[39] Lifton 1986, 106.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clues that betray the presence of aspects of themselves that needed to be suppressed in order for them to act like Nazis.

5.6 How to Kill Millions Without Noticing

We have seen that in order to be able to act wickedly, the Nazi physician had to keep parts of himself walled off. This psychic division enabled him to resist the operations of thoughts, principles and feelings that normally oppose genocidal behavior. This displays that the Nazi doctors tended to lacked integrity, that they were not willing to confront those aspects of themselves resisting their participation in the killing. It would be impossible, however, to find one particular characteristic shared by all Nazi doctors that displays that they lacked integrity, and that this lack of integrity was due to their failure to bring *their moral views* to bear on their lives. A more realistic strategy would be to take a more biographical approach to the matter, to examine the lives of representative individuals for the presence of moral views walled off from action.

The best example of a Nazi suffering from a lack of integrity may be the chief doctor at Auschwitz, Dr. Eduard Wirths. On many counts Wirths struck observers and even some of his Jewish patients as a decent man. On the other hand, Wirths bears the responsibility for the deaths of some four million human beings.[40] Wirths

[40] Lifton 1986, 384.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was no stranger to good deeds; early on he would treat Jewish patients after it had become illegal for Aryan doctors to do so. Many Jews injured in persecutions would sneak into Wirths' office at night for medical help. Nevertheless, Wirths appeared to accept fully the Nazi ideology about the threat Jews posed toward the German state. His ideology and his practice were at odds with one another: he could destroy "the Jew" while taking personal risks to help individuals he knew were Jewish.

At Auschwitz, Wirths' self fragmented still more. After another officer had selected two thousand Jewish patients for the gas chamber, some prisoner doctors pleaded with Wirths to spare them. The patients were healthy and could do good work. Wirths nearly in tears agreed to arrange for their safety. A few days later, however, Wirths himself selected two thousand other people in a different subcamp. He did not want others to see him participate in killing, nor did he think of himself as a killer. But, of course, he most certainly was.[41]

The most remarkable aspect of Wirths' doubling, however, concerns his own conception about his relation to the Jews. In his autobiography Wirths amazingly and self-servingly writes that "it probably can be credited to me that Jews are alive in Europe at all today."[42] For Wirths thought of himself not as a facilitator of genocide, but as a physician on a crusade to stamp out disease, to heal the prisoners, to improve medical conditions and supplies, and to lower the death rate in the camps. In his

---

[41] Lifton 1986, 395.

[42] Lifton 1986, 401.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

correspondence with family, he expressed his opposition to the mass killings that he in fact supervised. And indeed there are many documented cases of him doing out of the ordinary things to save individual Jewish prisoners. But any credit he deserves for such acts is obliterated by his active complicity with the goings-on at Auschwitz: Wirths damned himself by supporting Hitler's 1939 warning that the Jews would be destroyed should war break out, and by proudly claiming that he had "never done anything contrary to what was expected of me."[43]

As we saw to be the case for Nazi physicians generally, Wirths too conducted nightmarish "medical experiments" whose putative benefits to others he could distract himself with. His attitude toward selections was no different; Wirths thought that the physician needed to evaluate who was well enough to work, knowing full well the actual point of selections. As we saw earlier, Wirths intended to reduce the death rate at Auschwitz, but was willing to take the most contradictory means to that end.

After the war Wirths could admit that "[i]t was insane that people whom one has saved through the efforts and art of the physicians, who because of treatment, improvement in diet, personal hygiene, etc. were now forcibly killed"; but he never fully acknowledged that he was as responsible for the killing as anyone else.[44] Wirths intended to lower the death rate in Auschwitz; he was responsible for the deaths of millions. We might say that his physician self, which we might say 'contained' the

[43] Lifton 1986, 404.

[44] Quoted by Lifton 1986, 405-6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

intention to lower the death rate, was at odds with his Nazi-Auschwitz self, which 'contained' the intention to carry out orders. Obviously he could not do both, but Wirths' psyche was so fragmented that this fact did not prevent him from trying.

Though Wirths' doubling was more extreme that usual, Lifton nonetheless concludes that Wirths was all too representative of the physician's corruption in Nazi Germany.[45] Analogous biographical investigations into the lives of other Nazi physicians would likely uncover evidence of their psychic division, division that permitted and promoted wrongdoing despite the presence of moral views opposing their work.[46] The typical Nazi wrongdoer had aspects of himself that opposed what he in fact did, but these aspects were rendered inoperative.

I think we acknowledge this fact when we say things like "He knew better than to follow orders." We suspect that there are pockets of humanity somewhere buried in him who is otherwise a demon. As monstrous as most of his actions were, Wirths nevertheless had these pockets of humanity in him, and it is not a stretch for us to suspect that Wirths is far from alone in having moral views that need to be actively suppressed in order to act immorally.

---

[45] Lifton 1986, 414.

[46] See Levi 1989 [1958], 25-27, in which Levi concludes from the records of the justifications and rationalizations Nazis gave of their actions that they deceived themselves about what they were really doing.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.7 FACE-TO-FACE EXPERIENCES

We can get a better picture of the nature of the lack of integrity suffered by the Nazis by noticing what things brought about and fortified psychic fragmentation. As is well-known, Freud maintained that strong wishes cause certain kinds of irrationality, kinds akin to the ones we are concerned with here. Elements of the mind do not exert their normal rational influence upon one another, Freud thought, because it serves some psychic need to believe that things are as one wishes.

Now there is no doubt that much of the practical irrationality demonstrated in the lives of Nazi physicians was due to the work of unruly wishes. Eduard Wirths, for one, strongly believed that Hitler had to be ignorant of the mass killing projects that Wirths himself supervised. Wirths wanted to believe that the genocide would eventually stop, that a triumphant Nazi Germany would eventually shun such activity. It was perhaps the only way Wirths could continue to remain loyal to Hitler. But, of course, Wirths had no evidence of Hitler's ignorance; his attributing such ignorance to Hitler was a result of wishful thinking.

But any investigation of the Nazi physicians will lead to the observation that their fragmentation stemmed in large part, not only from unruly wishes, but also from the effects of a campaign of depersonalization that operated on both individual and organizational levels. The lack of personal interaction with potential and actual victims enabled their psychic fragmentation to occur and persist. Although it surely is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

an exaggeration, even Eichmann claimed that he refused to watch the killing of the Jews unless formally ordered to do so.[47] The Nazi's ability to comply with genocide largely depended on whether he avoided face-to-face experiences with those he was to wrong.

Even before killing became part of their job, the Nazi physicians were generally strongly anti-Semitic, an ideology which surely made it easier for them to adapt to their role. They could bring themselves to kill those whom they viewed as a threat. Problems often arose, however, whenever anti-Jewish activity was directed at individual Jews with whom they had personal contact. It was generally much more difficult for them to mistreat fellow Jewish doctors with whom they may have worked than for them to mistreat those whom they knew only as Jewish. When one Nazi physician found out that one of his prisoners had attended the same university as had he, his attitude to her improved remarkably: they would reminisce together about professors and shops, he gave her personal attention when she suffered from typhus, warned other SS personnel not to mistreat her, and so on.[48] Himmler famously complained that eighty million Germans each knew their own 'decent Jew' they wanted to save, the Jew whose existence wasn't a problem and so didn't require a solution.[49] While on the one hand Nazi doctors believed that the Jews were a danger

[47] Todorov 1996, 162.

[48] Lifton 1986, 231.

[49] Arendt 1994 [1962], 133.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to Germany and to the Aryan race, they also knew Jewish individuals who were neither of these things. Lifton concludes that

> [in] virtually all cases the Auschwitz self sought to block out potentially guilty images of actual Jews in favor of an ideological vision of constructive purpose in eliminating Jews or of 'solving' the 'Jewish problem.' There were conflicts in that combined stance, as we know, but mostly of a kind that did little to interrupt the work of the Auschwitz self.[50]

In this light, the ghettoization of the Jews can be seen to have served an ulterior purpose. By isolating Jews for Germans, it became much easier for Germans to swallow Nazi propaganda about the Jewish people. The fewer Jews a German personally knew, the more likely the Nazi picture of the Jew dominated the German's mind.

Depersonalizing the prisoners whom Nazi doctors *did* encounter was an active and ongoing project. The Nazi physicians "called forth every possible mechanism to *avoid taking in psychologically what they were doing*—every form of psychic numbing and derealization."[51] Primo Levi relates his experiences with a Nazi official who apparently was very good at not seeing individuals as such:

> Pannwitz is tall, thin, blond: he had eyes, hair and nose as all Germans ought to have them, and sits formidably behind a complicated writing-table. I, Häftling 174517, stand in his office, which is a real office, shining, clean and ordered, and I feel that I would leave a dirty stain whatever I touched.
>
> When he finished writing, he raised his eyes and looked at me.

---

[50] Lifton 1986, 438-9.

[51] Lifton 1986, 200.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> Because that look was not one between two men; and if I had known how completely to explain the nature of that look, which came as if across the glass window of an aquarium between two beings who live in different worlds, I would also have explained the essence of the great insanity of the third Germany.[52]

Levi not only bears witness to the thought that the Nazis had psychologically adapted to not seeing individuals as such, he also confirms our suspicion that understanding this feature of their psyche is crucial for our grasping their capacity for evil.

It is not an oversimplification to claim that the Nazis were socialized by their surrounding institutions to avoid seeing their victims as individuals, as persons. Todorov concludes that "all possible measures were taken in the concentration camps to ensure that face-to-face encounters did not occur, to prevent the executioner from meeting his victim's gaze."[53] Nazi officials were constantly looking for ways to enable their personnel to kill "life unworthy of life." A number of things contributed to their success.

As is well-known, Nazi Germany expended much effort to detoxify their language. One scholar of the Holocaust reports that after scouring through tens of thousands of Nazi documents, he never encountered the work "killing" except once in reference to an edict concerning dogs.[54] Euphemisms and code terms abounded for

---

[52] Levi 1960, 105.

[53] Todorov 1996, 161.

[54] Hilberg 1978, 275 as reported by Lifton 1986, 445.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

their unjust work, euphemisms that contributed to the sense of unreality, numbing, and "middle knowledge" about what they were doing.[55]

But language could also be the key for survival. Survivors report that the best way to get better treatment from the guards was to establish a personal relationship with one of them, and the only way one could do that is if one knew how to speak the same language. Without a common language, it was easier for the guards to ignore the prisoner's humanity.[56]

A vicious circle thus manifested itself in the relation between concentration camp officers and prisoners.[57] Prisoners were stripped of a name, forced to appear naked, huddle in masses, live in filth, scavenge for food; in short, they were forced to live like animals. And so it became easier for the officers and guards to ignore their humanity, to see them as animals, and so to impose further dehumanizing conditions upon them. One camp commander admitted, "I rarely saw them as individuals. It was always just a huge mass."[58] The point of all of this humiliation was to make it

---

[55] Hilberg 1985, 133-134, 167, 282. Lifton 1986, 445. Lifton (1973, 353) reports that the code names chosen for American weapon systems in the Vietnam War (e.g. Grasshopper, Comfy Bee, Puff the Magic Dragon) tended to sound more innocuous the more dangerous the weapon.

[56] Todorov 1996, 159.

[57] Todorov 1996, 160-161.

[58] Todorov 1996, 161.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

possible for the guards, officers, and others to carry out orders; otherwise, they would not have become the mass murderers they became.[59]

As discussed above, routinization also proved important and often necessary in order to keep one's humane thoughts isolated. For example, German railroad officials responsible for transporting Jews from ghettos to death camps found that they could avoid confronting their guilt by not varying their routine one bit.[60] Fixing the nature and the pattern of their work enabled them to avoid thinking about what they were doing, to avoid thinking about the consequences of their actions, really to avoid thinking entirely. These railroad officials *made themselves* cogs in a social machine. Much the same was also true of the work in the mobile killing operations and in the death camps.[61] Doctors working in the death camps fared better when the institution of selections replaced an environment where both camp officials and prisoners would simply beat weaker prisoners to death. Even though the hands of the doctors became dirtier, things became much more routinized, institutionalized, and so provided means for avoiding acknowledging the meaning of one's actions. This kind of social structure made it possible for individuals to shun taking personal responsibility for what they were doing.

---

[59] Levi 1989 [1958], 125-6.

[60] Hilberg 1978, 273. Cf. Hilberg 1985, 288, 292.

[61] Hilberg 1985, 126. Lifton 1986, 459.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Even the lives of top Nazi officials display this kind of evasion. In thinking about Albert Speer, Hitler's chief architect, Tzvetan Todorov dwells on the ways in which the Nazi emphasis on instrumental thought led to a "forgetting of ends." Speer explains that he was

> the top representative of a technocracy which had without compunction used all its know-how in an assault on humanity. . . . [B]y my abilities and my energies I had prolonged that war by many months. . . . Although I never actually agreed with Hitler, . . . I had nevertheless designed the buildings and produced the weapons which served his ends.[62]

Speer admitted that the sight of suffering people tugged upon his emotions, but his psyche was fragmented enough such that it had no influence upon what he did. His thoughts were completely absorbed with technical and instrumental questions. And what Speer could admit was true of himself is also no doubt true for many others in similar positions.

Perhaps the most striking indicator of the role of face-to-face encounters with one's victims is demonstrated by the evolution of the favored Nazi killing method. From 1939 to 1941, genocide was carried out mainly by the *Einsatzgruppen*, Nazi soldiers whose sole task it was to shoot Eastern European Jews face-to-face. Although very deadly, large numbers of the *Einsatzgruppen* nevertheless experienced severe psychological problems: anxiety, nightmares, hallucinations, tremors, deep guilt, and suicide.[63] Many of them had to be treated by Nazi psychiatrists, some had

---

[62] Speer 1970, 520-523. As quoted by Todorov 1996, 177.

[63] Hilberg 1985, 274.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to hospitalized. Alcoholism was actually encouraged; it was the only way most of the assassins could do their work.[64] Some tried to avoid these effects by employing massed fire from farther off, achieving psychological distance from what they were doing by putting physical distance between themselves and their victims.[65] Commanders did what they could to keep their methods as "humane" as possible, humane for the executioners rather than the victims; the commanders were aware of the weight of the psychic burden the riflemen were bearing.[66] After Himmler witnessed one of these mass killings—at which he was visibly shaken—he and other Nazi leaders came to recognize that this means of killing was not sustainable, that it was destroying the Germans as well as the Jews. Thus they searched for a less traumatic way to commit genocide.[67]

Experiments with carbon monoxide poisoning were tried for a while, but many victims exposed to CO did not die from the exposure, and so the Nazis still found themselves frequently using their firearms. Nearly as troublesome was the program of phenol injections physicians often performed upon individual prisoners.[68] This direct killing took its toll on all parties.

---

[64] Levi 1989 [1958], 31. Hilberg 1985, 146.

[65] Hilberg 1985, 126.

[66] Hilberg 1985, 127, 276-277.

[67] Hilberg 1985, 136-137, 274.

[68] Lifton 1986, 268.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It was not until the Nazis hit upon using Zyklon-B, a gas first used to kill rats, that it became possible to kill very large numbers of people without severe immediate psychological damage to the murderers. The victims were killed in closed chambers; no one had to watch them die. And even if they were watched, the victims became corpses without violent convulsion, without blood, without as much *visible* pain:

> Higher technologies render the killing more efficient, in time and numbers, and in easing the psychological burden of the perpetrators. A clear example is the Nazi sequence from face-to-face shooting to fatal injections and carbon monoxide gas chambers to cyanide gas: the sequence, to paraphrase Ernst B. [a Nazi Doctor], from pre-craftsmanship to craftsmanship to modern technology. The sequence helps eliminate the impediment of empathy, of experiencing one's victims as fellow human beings.[69]

The horrors of face-to-face killing had been greatly mitigated, though of course things were no less hellish for those killed.

This relationship between the capacity for killing and technological distance from one's victim is of course not limited to the Nazi experience. Psychological studies of American military in the Vietnam War reveal the ways in which technological interfaces between killer and victim enable the killer not to acknowledge fully what he is doing.[70] B-52 bomber pilots flying at very high altitudes could not see their victims on the ground, and later, in interviews about their work, they spoke only of their professional skill and performance. Pilots on lower-flying fighter-bomber missions could catch glimpses of people below, and felt some

---

[69] Lifton 1986, 493.

[70] Branfman 1971.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

need to rationalize what they did. They acknowledged that their activity was, though justifiable, problematic. Finally, those on low-flying helicopter gunships saw exactly what they were doing, and thus they tended to suffer most from severe emotional conflicts about the appropriateness of their activity.

It is unlikely that the soldiers in helicopter were generally more ethically sensitive than those in high-flying airplanes. Rather, the less one experienced one's victim's humanity, the easier it was to avoid thinking about the implications of one's actions. One had contact only with fellow soldiers and with one's machines. One's enemy was merely a blip on an electronic screen, and so one didn't even need to hold a racist ideology in order to kill them without conscious compunction. One could kill without noticing whether one's killing was consistent with the rest of what one believes and prizes. It was easier for one to lose one's integrity.

Similar conclusions were reached by Stanley Milgram in his famous psychological experiments. Although the primary purpose of the experiment was to determine whether a subject would obey an authority figure, even when this meant causing another serious pain, the findings also display that people are less likely to harm those with whom they have face-to-face contact:

> In these experiments, as the victim was brought closer to the man ordered to give him shocks, increasing numbers of subjects broke off the experiment, refusing to obey. The concrete, visible, and proximal presence of the victim

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> acted in an important way to counteract the [authority figure's] power and to generate disobedience.[71]

Milgram surmises that when one is in the presence of a potential victim, one's empathic responses are triggered, and one has a more complete grasp of the victim's experience. Otherwise, the victim's suffering can seem abstract and remote to the perpetrator. The perpetrator can put the victim out of mind, the effects of his actions do not seem to him tightly linked to his actions themselves.[72]

These and other technological interfaces make it easier for one to act in ways one otherwise would avoid. One doesn't *mind* so much the unfavorable consequences of one's actions. And so the way is cleared for one to do things one would balk at if one were to give the matter one's full attention. Of course, one often acts in ways that have unforeseen consequences. This in and of itself does not display that one lacks integrity. Rather, what matters is whether one fails to think about the consequences of one's action because at some level one surmises that a proper consideration of their importance would dissuade one from so acting. And certain kinds of technological interfaces between agent and patient have the capacity for promoting this sort of failure.

Technology, routinization, diction, and institutionalization all serve to depersonalize those whom one wrongs, and thus to avoid thinking about the

---

[71] Milgram 1974, 4.

[72] Milgram 1974, 36-39.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

implications of one's actions. When one fails to think about what one is doing, one is more likely to do things that conflict with one's other plans, projects, and principles. By avoiding face-to-face encounters with one's potential victims, one can keep buried aspects of oneself that would otherwise prevent one from actually victimizing these people, aspects such as the thought that one is not to treat human beings as mere things. And that is just what the Nazi totalitarian system did so effectively, making the killing process as mechanized as possible, relieving murderers of the need for the sort of deliberation that is characteristic of one with integrity.

The avoidance of personal encounters with their victims, then, enabled the Nazis to shut out aspects of themselves opposed to their work. Had they any thoughts about treating Jews like human beings—which many of them did—the campaign of depersonalization totalitarianism promotes so well pushed these into the background, making sure that they never really took the trouble to act upon them. The Nazi's lack of integrity was sustained by the fact that he did not encounter his victim in a face-to-face manner. The lack of face-to-face experiences with those with whom one deals can explain why one (mis)treats them in ways one wouldn't if one were being honest with oneself.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

5.8 Drawing more general lessons

Thankfully, very few human beings are mass murderers. (How could there be many?) Fewer still are guilty of genocide. Yet if we can begin to understand how many fairly ordinary, educated, cultured moderns acted monstrously, we will be in a better position to understand more common forms of immorality. Specifically, grasping how the avoidance of face-to-face encounters with those whom we treat wrongly enables us to hold at bay aspects of ourselves that would resist our acting uncaringly—grasping *this* takes us a considerable way toward explaining why we act immorally.[73] Not that all immorality can be chalked up toward avoiding acknowledging the humanity and the personality of the other. All too often we wrong those closest to us.

But much of our wrongdoing stems from avoiding the gaze of the other, from seeing individuals only as members of categories, as abstractions.[74] We care about the well-being of others, but much of the time we block out this part of ourselves. Thinking about how we fail to think about other people draws our attention to the fractures that run through our lives. By not fully attending to those with whom we affect, we often end up wronging them, suggesting that our immorality can frequently

[73] Cf. Hilberg 1985, 263.

[74] At the funeral of Emmanuel Lévinas—a philosopher much concerned with the face of the Other—Israeli philosopher Zev Harvey reported that Yitzhak Rabin, right after he was shot, turned around to look into the face of his killer. One witness reported that the killer looked to the other side, could not look Rabin back.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

be explained by noting—not only our failure to hold the correct moral views—but our lack of integrity, our failure to act in light of what we deep down know to be true.

However, to claim that all immoral action stems from a lack of integrity would be dubious, surely an overgeneralization. But the case for thinking that immorality is, in one way or another, a product of error is looking stronger. As we saw earlier, the resources available for explaining the incorrectness of the immoral person's views about what he ought to do are considerable, though in the end admittedly difficult to assess. Now we can also see that there are grounds for thinking that many cases of immorality flow not from ignorance about right and wrong, but from agents who lack integrity, agents who suppress their moral views such that their actions are largely untouched by their deep attitudes about what they should do. Moral truths, then, can fail to affect an agent's actions either because he doesn't know about them or because he doesn't act in light of them.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Conclusion

Rather than simply restating conclusions I have already drawn, perhaps it would be more instructive to relate what I have done here to some traditional work in moral philosophy. Briefly put, I have employed broadly Aristotelian resources to address broadly Socratic problems.

In the early Platonic dialogues, Socrates argues that knowledge of the good is sufficient for virtue (*Laches*, 199de; *Gorgias* 488a). He holds that those who lack virtue also lack this knowledge. Immorality is ultimately a cognitive defect, and so the road to virtue must lead through enlightenment. But he also thinks that no one in fact possesses knowledge of the good, and thus that each of us lacks virtue.[1] Furthermore, he not only believes that each of us is ignorant of the good; he also thinks it important to demonstrate that each person lacks this knowledge. He is not content merely with asserting that others are in error.

---

[1] There is serious debate about whether Socrates' profession of ignorance is sincere or ironic. This is not the place for me to enter this debate, but, for what it is worth, let me say that I suspect that Socrates thinks that he has true beliefs about the good, even though he cannot prove that they are true to his satisfaction, and thus he concludes that he lacks *knowledge*. On this view, Socrates is being truthful when he claims that he and everyone else lacks knowledge.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

How does Socrates aim to show that others are in error? The particular method he adopts to reveal that his interlocutor lacks knowledge of the good, the *elenchus*, aims to demonstrate that his interlocutor holds inconsistent beliefs about the good. The pattern of the elenchus, reconstructed as follows by Gregory Vlastos[2], is:

(1) The interlocutor asserts a thesis, *p*, which Socrates considers false and targets for refutation.

(2) Socrates secures agreement to further premises, say *q* and *r* (each of which may stand for a conjunct of propositions). The agreement is *ad hoc*: Socrates argues from {*q*, *r*}, not to them.

(3) Socrates then argues, and the interlocutor agrees, that *q* & *r* entail *not-p*.

(4) Socrates then claims that he has shown that *not-p* is true, *p* false.

Socrates shows that there is an inconsistency in the belief-set of the interlocutor, and that one of the theses in questions must be rejected. Further, Socrates seems confident that he can show that this is so to *any* willing participant who holds moral views at odds with his. His confidence is justified because he thinks that every person deep down *already* believes *not-p*, or at least believes other theses that in fact entail not-*p*, even if that person explicitly insists that *p*. And so employment of the elenchus will reveal her inconsistency if she continues to assert *p*. On this view, people already have true moral beliefs; they assert otherwise only because their souls are in some sense divided or fragmented.

---

[2] Vlastos 1994, 11. I will not rely upon the controversial fourth step.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

So Socrates thinks that:

(S1) The not-fully-virtuous person lacks knowledge.
(S2) The not-fully-virtuous person is at odds with himself.
(S3) It is important to demonstrate, rather than merely assert, S1 and S2.

I have been trying to make it seem more plausible that, for any immoral person, *either* S1 *or* S2 is true.[3] These were the goals of Chapters Four and Five, respectively. And in Chapters Two and Three, I essentially argued for a version of S3.

But the strategies I have favored for securing these theses differ radically from those adopted by Socrates. In fact, the strategies I have favored owe more to the work of Aristotle. This should be surprising, for Aristotle himself does not unambiguously hold S1-S3.

Although Aristotle often says things which indicate that he roughly agrees with S1[4], he also says things that seem to suggest otherwise.[5] He rejects S2; on his view, the vicious person need not suffer any internal division.[6] And even if Aristotle were to accept S1 and S2, he would still reject S3, for he is not concerned to demonstrate that the immoral person is making some mistake or other, not as Socrates is. On the contrary, Aristotle does not seriously consider the views of those who exclusively pursue base things, the views of those whom he judges to be badly

---

[3] I am not at all confident that for all immoral persons *both* S1 and S2 are true.

[4] Roughly, because Aristotle distinguishes various intellectual virtues (*episteme*, *sophia*, *phronesis*, *gnome*, *sunesis*, *euboulia*). See *Nichomachean Ethics* VI. 12 (1144a35-36).

[5] For example, see the difficult *Nichomachean Ethics* VII. 2-3.

[6] See *Nichomachean Ethics* VII. 1; VII. 8.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

brought up; they are not proper interlocutors.[7] He seems far less concerned than Socrates to ascertain that such people *really are* in a poor position to know how to live.

But on the other hand, Aristotle does stress the importance of listening to *endoxa*, the reputable opinions of the wise and the many. Aristotle even regards the fact that *p* agrees with endoxa to count as evidence for *p*'s truth. Socrates rarely makes that inference; rather, Socrates is concerned with whether only *his interlocutor* accepts *p*, can explain why *p*, can hold *p* and all his other's beliefs consistently.[8] Aristotle is inclined, as Socrates is not, to count reputable testimony as a way to reasonably acquire moral views. On his view, there are occasions in which it is appropriate to trust what another person says about how to live.

Another relevant point of contrast is that while Socrates seems to think that we become good through philosophical argument, Aristotle holds that we become good (also) through the repeated performance of good actions, through our experiences. Only those who have undergone the right kinds of experiences know enough to be proper listeners to philosophical discussion about the good life; only those who have undergone the right kinds of experiences make for proper advisors. The importance of one's past pervades key aspects of Aristotle's moral philosophy.

---

[7] Aristotle, *Nichomachean Ethics*, I.3 (1095a1-13), I.4 (1095a28-30)

[8] Vlastos 1994.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Further, Aristotle thinks the incoherences that *do* run through the life of the not-fully-virtuous person are revealed not by getting her to state views that can be shown to be contradictory. Or at least this is not the only means for revealing them. Rather, Aristotle thinks that a person's *actions* can be in conflict with her views, that mental conflict is often revealed by her saying one thing and yet doing another.

So Aristotle, unlike Socrates, thinks that:

(A1) It can be reasonable to trust another person's advice about what is important.
(A2) A person becomes good in part through experience.
(A3) The fragmented nature of a person's soul is largeiy revealed by noting the conflicts between her actions and her views.

Now I also have argued for A1-A3, and, furthermore, I have argued that the truth of A1-A3 helps us make a better case for S1-S3. The Aristotelian theses strengthen the bold Socratic claims.

In Chapter One, I have argued that a rational agent trusts the advice of those whose advice is trustworthy (A1). The rational agent does not always act in ways that accord with her own practical views; sometimes she sets aside her own view about what to do because she thinks that another may be better suited to decide. However, the rational agent does not give up her rationality when she trusts the advice of another; on the contrary, she exercises her rationality in deciding whom to trust and when to trust him. Part of being rational involves putting available resources to use. One way to fail to act rationally is, through stubbornness or hubris, to ignore advice worthy of trust.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I also have argued that the an agent's actions are rational to the degree that she acts in the light of available (hers or others') life-experiences (A2). Just as our conclusions can be criticized either if we are ill-informed, or if we ignore available information in reaching those conclusions, so too our actions can be criticized if we are inexperienced, or if we ignore lessons learned from available life-experiences in deliberation. The rational agent is not the agent who is completely innocent of the ways of the world, naïveté is neither a theoretical nor a practical virtue.[9] Of course, accepting this does not necessarily require us to deliberately seek out and undergo experiences that we suspect will be educative. Whether to do this or not is *itself* a practical question, one which various people may be better or worse suited to answer. I hope to have shown merely that it is a live question.

Fortunately, sometimes we can bypass acquiring certain educative life-experiences for ourselves by listening to those who have already undergone them. Often we trust the advice of others, and when we search for the reason that justifies our deference, we find that the experience of others is a better answer than many of the other options. And once we see that our experiences can be educative, we will probably find it appropriate to trust the advice of others in more situations than we otherwise might. A1 and A2 thus combine in ways that enrich our vision of rational agency.

---

[9] Cf. Hampshire 1989.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As I argued in Chapter Four, A1 and A2 together also make S1 seem more plausible: once we recognize that our inexperience impairs our own views about how to live, we have stronger grounds for thinking that those who are immoral also probably lack the kinds of experience for knowing how to live. And so questions surrounding the objectivity of morality might be more fruitfully addressed if we bear in mind the ways that we are agents both who reasonably trust the advice of others and whose life-experiences improve our practical thought; for insofar as these questions concern figuring out whether deviations from the moral life are thereby mistakes, we can better answer them by reflecting upon the degree to which immorality stems from a lack of experience.

As I argued in Chapter Five, A2 and A3 make S2 seem more plausible than it otherwise might. The lack of certain kinds of life-experience can make it more difficult for us to put our views about how to live into practice. Sometimes it is not enough to possess correct views about what to do; in order for them to govern our conduct effectively, we need to encounter those we value in the proper light. Otherwise, we act in ways contrary to what we deep down think is best; we suffer from a lack of integrity.

My attitude toward S3, however, is neither as enthusiastic as Socrates', nor as uninterested as Aristotle's. I don't think it is important to demonstrate *to* each not-fully-virtuous person that either she lacks knowledge or that she is at conflict with herself. This was the result of Chapter Three. But it is important to have detailed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

explanations for why the immoral are in error; it is not enough merely to assert that they are. This was the result of Chapter Two.

But have I in fact shown that, for any immoral person, either S1 or S2 is true? Have I cleared away (what I take to be) the most significant obstacle to embracing the objectivity of morality?

Unfortunately, there may still exist the immoral agent who lacks no experience helpful either for grasping practical truths or for translating those truths into action. We have no grounds for presuming that such a character is only fictional. Confronted with such a person, it seems that we would have to conclude that claims to moral objectivity are bluff, that while we can rightly regard him as wicked, evil and callous, we cannot rightly regard him as mistaken or irrational.

Thus it is possible that morality is not objective. Whether this is so, however, depends not only upon more philosophical argument, but upon quasi-empirical investigation as well. Our moral views are nourished not only by information and by armchair reflection on the ways in our philosophical theories cohere with these views. They are aided also by the experience of history, autobiography, social psychology, and especially life itself.[10] I hope to have illustrated some of the types of investigation that would prove helpful, both with the discussion of the ways in which the lack of various kinds of educative experiences impairs our practical judgment, as well as with

[10] See Nussbaum 1990, 45-49, with which I differ only on particular emphasis.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the discussion of how one group of the immoral suppressed their moral views in order to murder. Obviously, there are many other facets of the topic to be investigated, some which may make the case for moral objectivity seem stronger, as those I have focused upon here, others which may not. There are plenty of opportunities for further work.

We should also consider more carefully the objectivity of different parts of morality. That is, we might find that we can explain some but not all types of immorality as products of error. There is no *a priori* reason not to think that some types of moral truths will be objective while others are not; it depends rather on the resources available for showing that various deviations from a moral life result either from ignorance or from a lack of integrity. Our resources for explaining why the wicked are in error are potentially great, but it is unclear whether they are in fact enough.

Prospects look good, I think, for showing that the perpetrators of extreme forms of violent injustice are in error: we can show that they are ill-informed, lack the right kind of experiences with those whom they wrong, and/or are self-deceived. This is what our examination of the Nazi physicians and killers strongly suggests. And if executive virtues such as courage and temperance are part of the content of morality, then it also probably can be argued that if one has integrity, one has these moral virtues as well, though I won't do that here.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Trouble arises, I suspect, when we turn to the free rider—the person who benefits from collective action in which he refuses to participate—and the promise-breaker—the person who fails to keep his promise when it is in his interest to break it. It just seems too plausible that there are human beings who violate their word and mooch off of others who neither lack the sort of experiences that provide practical knowledge, nor are in conflict with themselves about what they do. But even here I think the moral objectivist can make a case.

We may be able to argue that the experienced person would have a *disposition* to keep promises and cooperate with others, and that one can have integrity only if one's dispositions sit deep, carrying one through the tight corner and the opportunity to freeride. But thinking about the ways in which the rationality of actions depends upon an agent's dispositions is a topic we have to leave for another time.[11]

[11] See Thompson 1992.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bibliography

Altham, J.E.J. and Ross Harrison, eds. 1995. *World, Mind, and Ethics: Essays on the ethical philosophy of Bernard Williams.* Cambridge: Cambridge University Press.

Anderson, Elizabeth. 1993. *Value in Ethics and Economics.* Cambridge, Massachusetts: Harvard University Press.

Anscombe, G.E.M. 1957. *Intention.* Ithaca: Cornell University Press.

——. 1981. *The Collected Philosophical Papers of G.E.M. Anscombe*, vols. 1-3. Minneapolis: University of Minnesota Press.

——. 1982. "Medalist's Address: Action, Intention and 'Double Effect'." *Proceedings of the American Catholic Philosophical Association* 56.

Arendt, Hannah. 1994 [1962]. *Eichmann in Jerusalem*, rev. ed. New York: Penguin Books.

Aristotle. 1982. *Nichomachean Ethics*, trans. Ross and Urmson, in *Complete Works*, ed. Barnes. Princeton: Princeton University Press.

Blackburn, Simon. 1984. *Spreading the Word.* Oxford: Oxford University Press.

Branfman, Fred. 1971. "The Wild Blue Yonder Over Laos." *Washington Monthly* (July).

Burnyeat, M. F. 1980. "Learning to be Good." In Rorty, 1980.

Cartwright, Nancy. 1983. *How the Laws of Physics Lie.* Oxford: Clarendon Press.

Coady, C.A.J. 1992. *Testimony: A Philosophical Study.* Oxford: Oxford University Press.

Cohon, Rachel. 1986. "Are External Reasons Impossible?" *Ethics* 96: 545-556.

Copp, David. 1990. "Explanation and Justification in Ethics." *Ethics* 100.

Craig, Edward. 1990. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis.* Oxford: Oxford University Press.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Daniels, Norman. 1979. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *The Journal of Philosophy*. 76: 256-282.

Darwall, Stephen. 1983. *Impartial Reason*. Ithaca: Cornell University Press.

Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.

Descartes, Rene. 1985 [1637]. *Discourse on the Method* in *The Philosophical Writings of Descartes*, trans. Cottingham, Stoothoff and Murdoch, vol.1. Cambridge: Cambridge University Press.

Diamond, Cora. 1991. *The Realistic Spirit*. Cambridge, Massachusetts: MIT Press.

Dummett, Michael. 1978. *Truth and Other Enigmas*. London: Duckworth.

Eggerman, Richard W. 1979. "Moral Philosophers and Moral Advisers." *Metaphilosophy* 10: 161-168.

Foley, Barbara. 1982. "Fact, Fiction, Fascism: Testimony and Mimesis in Holocaust Narratives." *Comparative Literature* 34: 330-360.

Foot, Philippa. 1994. "Rationality and Virtue", in *Norms, Values and Society*, ed. H. Pauer-Studer. The Netherlands: Kluwer Academic Publishers.

Frady, Marshall. 1996. *Jesse: The Life and Pilgrimage of Jesse Jackson*. New York: Random House.

Frankfurt, Harry. 1988. *The importance of what we care about*. Cambridge: Cambridge University Press.

Fussell, Paul. 1989. *Wartime: Understanding and Behavior in the Second World War*. Oxford: Oxford University Press.

——. 1996. *Doing Battle: The Making of a Skeptic*. Boston: Little, Brown.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge, Massachusetts: Harvard University Press.

Geach, Peter. 1956. "Good and Evil." *Analysis* 17: 33-42.

Geuss, Raymond. 1981. *The Idea of a Critical Theory: Habermas and the Frankfurt School*. Cambridge: Cambridge University Press.

Goldhagen, Daniel Jonah. 1996. *Hitler's Willing Executioners*. New York: Knopf.

Grudin, Robert. 1990. *The Grace of Great Things*. New York: Ticknor & Fields.

Habermas, Jurgen. 1990. *Moral Consciousness and Communicative Action*, trans. Lenhardt and Nicholsen. Cambridge, Massachusetts: MIT Press.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Hampshire, Stuart. 1989. *Innocence and Experience*. Cambridge, Massachusetts: Harvard University Press.

Hare, R.M. 1975. "Rawls' Theory of Justice", in *Reading Rawls*, ed. Norman Daniels. New York: Basic Books.

Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.

——. 1977. *The Nature of Morality*. Oxford: Oxford University Press.

——. 1986. "Moral Explanations of Natural Facts--Can Moral Claims Be Tested Against Reality?" *The Southern Journal of Philosophy* 24, supplement.

Hilberg, Raul. 1978. "Confronting the Moral Implications of the Holocaust." *Social Education* 42.

——. 1985. *The Destruction of the European Jews*. New York: Holmes & Meier.

Hsu, Andrew. Manuscript. "On the Natural History of some Artifacts."

Hume, David. 1978 [1739]. *A Treatise of Human Nature*, 2nd ed., ed. Selby-Bigge. Oxford: Clarendon Press.

Hurley, Susan. 1989. *Natural Reasons*. Oxford: Oxford University Press.

Kolnai, Aurel. 1978. *Ethics, Value and Reality*. London: Athlone Press.

Korsgaard, Christine. 1986. "Skepticism about Practical Reason." *The Journal of Philosophy* 83: 5-25.

——. 1996. *The Sources of Normativity* (Cambridge, Cambridge University Press).

Lear, Jonathan. 1984. "Moral Objectivity", in *Objectivity and Cultural Divergence*, ed. S.C. Brown. Cambridge: Cambridge University Press.

Levi, Primo. 1989 [1986]. *The Drowned and the Saved*, trans. Raymond Rosenthal. New York: Vintage Books.

——. 1993 [1958]. *Survival in Auschwitz* [*If This is a Man*], trans. Stuart Woolf. New York: Collier Books.

Lifton, Robert Jay. 1973. *Home from the War*. London: Wildwood House.

——. 1986. *The Nazi Doctors: Medical Killing and the Psychology of Genocide*. New York: Basic Books.

Mackie, J. L. 1976. *Problems from Locke*. Oxford: Clarendon Press.

——. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin Books.

McDowell, John. 1979. "Virtue and Reason." *The Monist* 62: 331-350.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

——. 1983. "Aesthetic value, objectivity, and the fabric of the world", in *Pleasure, Preference, and Value: Studies in Philosophical Aesthetics*, ed. Eva Schaper. Cambridge: Cambridge University Press.

——. 1985. "Values and Secondary Qualities", in *Morality and Objectivity: A Tribute to J. L. Mackie*, ed. T. Honderich. London: Routledge and Kegan Paul.

——. 1995. "Might there be external reasons?" In Altham and Harrison, 1995.

Milgram, Stanley. 1974. *Obedience to Authority*. New York: Harper & Row.

Miller, Richard. 1985. "Ways of Moral Learning." *Philosophical Review* 94: 507-556.

Millgram, Elijah. 1997. *Practical Induction*. Cambridge, Massachusetts: Harvard University Press.

Murdoch, Iris. 1956. "Vision and Choice." *Proceedings of the Aristotelian Society*. supp. vol. 30: 32-58.

Nagel, Thomas. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.

——. 1980. "The Limits of Objectivity", in *The Tanner Lectures*. Salt Lake City: University of Utah Press.

——. 1986. *The View from Nowhere*. Oxford: Oxford University Press.

Neely, Wright. 1974. "Freedom and Desire." *Philosophical Review* 83.

Nietzsche, Friedrich. 1966 [1886]. *Beyond Good and Evil*, trans. Walter Kaufmann. New York: Random House.

Nisbet R. and L. Ross. 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewoods Cliffs, NJ: Prentice Hall.

Norman, Richard. 1997. "Making Sense of Moral Realism." *Philosophical Investigations* 20: 117-135.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, Massachusetts: Belknap.

Nussbaum, Martha. 1990. *Love's Knowledge*. Oxford: Oxford University Press.

——. 1995. "Aristotle on human nature and the foundation of ethics", in Altham and Harrison, 1995.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.

Pears, David. 1984. *Motivated Irrationality*. Oxford: Clarendon Press.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Quinn, Warren. 1993. *Morality and Action*. Cambridge: Cambridge University Press.

Railton, Peter. 1984. "Moral Realism." *Philosophical Review* 95.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, Massachusetts: Belknap.

——. 1975. "The Independence of Moral Theory from Metaphysics." *Proceedings of the American Philosophical Association* 48: 5-22.

Rorty, Amelie, ed. 1980. *Essays on Aristotle's Ethics*. Berkeley: University of California.

Rosati, Connie. 1996. "Internalism and the Good for a Person." *Ethics* 106: 297-326.

Ryle, Gilbert. 1949. *The Concept of Mind*. London: Hutchinson.

Smith, Michael. 1994. *The Moral Problem*. Oxford: Basil Blackwell.

Speer, Albert. 1970. *Inside the Third Reich*, trans. Richard Winston and Clara Winston. New York: Collier-Macmillan.

Stern, Robert. 1992. "The relation between moral theory and metaphysics." *Proceedings of the Aristotelian Society*. 90: 143-159.

Stewart, J. Douglas. 1978. "An Analysis of 'Advising'." *Educational Theory* 28: 202-213.

Sturgeon, Nicholas. 1992. "Nonmoral Explanations." *Philosophical Perspectives* 6.

Taylor, Charles. 1985. *Human Agency and Language*. Cambridge: Cambridge University Press.

Taylor, Gabriele. 1981. "Integrity." *Proceedings of the Aristotelian Society*, supp. vol. 55: 143-159.

Thomas, Laurence. 1993. *Vessels of Evil: American Slavery and the Holocaust*. Philadelphia: Temple University Press.

Thompson, Michael. 1992. *Some Remarks on the Role of Generality in Practice*. Doctoral dissertation, University of California, Los Angeles.

——. 1995. "The Representation of Life", in *Virtues and Reasons: Philippa Foot and Moral Theory*, eds. Hursthouse, Lawrence and Quinn, pp. 247-296. Oxford: Clarendon Press.

Todorov, Tzvetan. 1996. *Facing the Extreme: Moral Life in the Concentration Camps*, trans. Abigail Pollak and Arthur Denner. New York: Henry Holt.

Vlastos, Gregory. 1994. *Socratic Studies*, ed. Myles Burnyeat. Cambridge: Cambridge University Press.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Wertheimer, Roger. 1971. "Understanding the Abortion Argument." *Philosophy and Public Affairs* 1: 67-95.

Wiesel, Elie. 1978. *A Jew Today*, trans. Marion Wiesel. New York: Random House.

Wiggins, David. 1976. "Deliberation and Practical Reason." In Rorty, 1980.

Williams, Bernard. 1972. "Knowledge and Reasons", in *Problems in the Theory of Knowledge*, ed. von Wright, pp. 1-11. The Hague: Martinus Nijhoff.

——. 1978. *Descartes: The Project of Pure Enquiry*. Harmondsworth: Penguin Books.

——. 1981. *Moral Luck*. Cambridge: Cambridge University Press.

——. 1985. *Ethics and the Limits of Philosophy*. Cambridge, Massachusetts: Harvard University Press.

——. 1995a. *Making sense of humanity*. Cambridge: Cambridge University Press.

——. 1995b. "Replies." In Altham and Harrison, 1995, pp. 185-224.

——. 1995c. "Ethics", in *Philosophy: a guide through the subject*, ed. A. C. Grayling, pp. 545-582. Oxford: Oxford University Press.

Wright, Crispin. 1992. *Truth and Objectivity*. Cambridge, Massachusetts: Harvard University Press.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.