

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Perfectionism, value pluralism, and the human good

Permalink

<https://escholarship.org/uc/item/05w5h7mb>

Author

Stedman, Jeffrey N.

Publication Date

2006

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Perfectionism, Value Pluralism, and the Human Good

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Philosophy

by

Jeffrey N. Stedman

Committee in charge:

Professor Richard Arneson, Co-chair
Professor David Brink, Co-Chair
Professor Alan Houston
Professor Richard Madsen
Professor Donald Rutherford

2006

Copyright

Jeffrey N. Stedman, 2006

All rights reserved.

The dissertation of Jeffrey N. Stedman is
approved, and it is acceptable in quality and
form for publication on microfilm:

Co-chair

Co-chair

University of California, San Diego

2006

DEDICATION

For Adam

TABLE OF CONTENTS

Signature Page.....	iii.
Dedication.....	iv.
Table of Contents.....	v.
Vita.....	vi.
Abstract.....	ix.
Chapter 1 Toward a Plausible Account of the Human Good.....	1
Chapter 2 Subjective Theories of the Good.....	15
Chapter 3 Hurka s Perfectionism.....	52
Chapter 4 Self-Realization Accounts of the Good.....	97
Chapter 5 Value Pluralism and Ethical Judgment.....	173
Bibliography.....	264

VITA

Areas of Specialization: Ethical Theory, Social and Political Philosophy, Philosophy of Law

Areas of Competence: Applied Ethics, History of Ethics, Metaphysics, Philosophy of Mind

Education:

University of Memphis

B.A. in Philosophy August, 1993, *summa cum laude*

Vanderbilt University Law School

Attended 1995-96 academic year

University of California, San Diego (Department of Philosophy)

M.A. in Philosophy March, 2003

Ph.D. in Philosophy September, 2006

Dissertation: *Perfectionism, Value Pluralism, and the Human Good*

Dissertation Committee: Richard Arneson (Co-chair), David Brink (Co-chair), Donald Rutherford, Alan Houston, Richard Madsen

Presentations:

Some Problems with Hurka's Perfectionism

University of California, San Diego Graduate Philosophy Colloquium, May 4, 2004

Comment on Eric Moore's Objective Consequentialism, Right Action, and Good People

American Philosophical Association, Pacific Division Meeting, March 23, 2005

Two Conceptions of Self-Realization

Alabama Philosophical Society, University of Montevallo (AL), October 21, 2005

Value Pluralism and Moral Judgment

Southern California Philosophy Conference, California State University at Northridge, October 29, 2005

Awards and Fellowships:

University of California, San Diego
Dissertation Fellowships, 2003-04
Research Fellowships, 2002, 2003

University of Memphis
Carol Bowman Scholarship (for outstanding Philosophy major), 1992

Teaching Experience:

As Instructor:

Introduction to Philosophy: Metaphysics; University of California, San Diego; Winter 2005
Medical Ethics; Christian Brothers University, Memphis, TN; Spring 2006
Elementary Logic; University of Memphis; Spring 2006

As Graduate Teaching Assistant:

Philosophy Department:

Introduction to Logic (Augustin Rayo, Araceli Penafuerte twice)
Logic and Decision Making (Jonathan Cohen)
Introduction to Philosophy: Ethics (Richard Arneson twice)
Ethics and Society (Mary Devereaux)
History of Philosophy: Philosophy in the Age of Enlightenment (Jay Odenbaugh)
Contemporary Moral Issues (Richard Arneson)
Biomedical Ethics (Gerald Doppelt, Jay Odenbaugh)
Philosophy of Law (David Brink)

UCSD s Revelle College Writing Program:

Humanities 1: The Foundations of Western Civilization: Israel and Greece
Humanities 2: Rome, Christianity, and the Middle Ages

Graduate Seminars:

Philosophy of Mind (*Psycho-neural Reduction*, Paul Churchland)
Ethics (*Ethics and Personal Identity*, David Brink)
Early Modern Philosophy (*Spinoza*, Nicholas Jolley)
Eighteenth Century Philosophy (*Free Will in Early Modern Philosophy*, Gideon Yaffe)
Contemporary Topics in the Philosophy of Science (*Science and Values*, Philip Kitcher)
Contemporary Epistemology and Metaphysics (*Quine*, Gila Sher)
Greek Philosophy (*Relativism and the Good in Ancient Philosophy*; Georgios Anagnastopoulos)
Ethics (*Contemporary Metaethics*, David Brink)

Philosophy of Mind (*Philosophy of Mind Since 1950*, Paul Churchland)
Social and Political Philosophy (*Rawls and Nozick*, Richard Arneson)
Ethics (*Contemporary Ethical Rationalism*, Richard Arneson)
Early Modern Philosophy (*Hobbes*, Donald Rutherford)
Ethics (*Analytical Jurisprudence*, David Brink)

Profession Memberships and Service:

Member of American Philosophical Association, 2000-Present
Referee for *Philosophical Psychology*, 2003

References:

David Brink
Professor of Philosophy
University of California, San Diego
Email: dbrink@ucsd.edu

Richard Arneson
Professor of Philosophy
University of California, San Diego
Email: rarneson@ucsd.edu

Donald Rutherford
Professor of Philosophy
University of California, San Diego
Email: drutherford@ucsd.edu

ABSTRACT OF THE DISSERTATION

Perfectionism, Value Pluralism, and the Human Good

by

Jeffrey N. Stedman

Doctor of Philosophy in Philosophy

University of California, San Diego, 2006

Professor Richard Arneson, Co-chair

Professor David Brink, Co-chair

What makes for a good life? If you have a child, a spouse, or a very close friend you probably want what is best for her. However, in order to know what would be best for your loved one you need an account of the personal good. I argue for a version of value pluralism, according to which there is disparate list of fundamental goods which resist reduction to some single supervalue such as pleasure and lack any strong unifying principle. I argue that any adequate account of the human good must include pleasure and freedom from pain, knowledge and understanding, practical rationality, and proper emotional responsiveness as fundamental goods, and that the only thing these various goods have in common is their intrinsic goodness. Value pluralism is an alternative to various kinds of monism about the good, such as hedonism, perfectionism, and the desire-satisfaction theory. Hedonism explains the value of any putative good in terms of pleasure, while perfectionism sees the good in

terms of the development and exercise of those capacities or characteristics essential to some aspect of our nature. Desire-satisfaction views are best considered a form of monism about the good, since their identification of one's good with the satisfaction of one's desires offers a way of unifying various goods.

The battle between monistic and pluralistic accounts is usefully analyzed in terms of competing theoretical virtues such as unity, simplicity, explanatory coherence, and plausibility. Monistic accounts typically score high along the first three of these dimensions and quite low along the last. I consider what I take to be the most promising versions of value monism, and argue that each scores so low along the dimension of plausibility that it should be rejected in favor of value pluralism. I begin by considering hedonism and the desire-satisfaction view and then discuss two distinct versions of perfectionism. Hedonism and perfectionism, in particular, are implausible because they fail to recognize our complex natures as embodied rational agents, and our need for a vision of the good which is correspondingly complex. In the final chapter I offer just such an account.

Chapter 1 Toward a Plausible Account of the Human Good

Theories of the good hold a central place in our systematic thinking about morality. This is obviously the case for consequentialists: if we are to maximize or otherwise promote the good, we must have some idea of what the good consists in. But even many deontologists, while they would deny that right action is simply a matter of promoting the good, would nevertheless agree that we have something like a prima facie or pro tanto duty to promote the good, though this duty may be overridden by competing considerations.¹ Moreover, one of the most prominent of contemporary deontologists, John Rawls, says that all ethical doctrines worth our attention take consequences into account in judging rightness. One which did not would simply be irrational, crazy.² Clearly, we need some kind of systematic account of the good. But to what extent can we hope to develop an account of the good which combines familiar theoretical virtues such as simplicity, unity, and determinateness with the potentially competing virtues of plausibility and livability? I believe that many traditional monistic theories of the good such as hedonism and perfectionism pay too high a price in plausibility and livability, such that allegiance to them consists mainly in the base coin of lip

¹ W. D. Ross uses the language of prima facie duties. It has been pointed out that this is misleading, since it suggests an epistemological criterion and Ross seems to have meant something quite different (see, e.g., Gaut, *Justifying Moral Pluralism* 138). Thus many prefer to substitute pro tanto for prima facie. In *A Theory of Justice*, Rawls defines deontological moral theories as those that either define the right independently of the good or deny that the right act is always whatever would maximize the good (30). On this understanding, Ross's view, along with many others, counts as a version of deontology, even though Ross thinks we have a prima facie or pro tanto duty to promote the good. A Nozickian-inspired view holding that we ought to promote the good insofar as this would not violate certain side constraints in the form of individual rights would also count as a form of deontology (see Nozick 28-33).

² Rawls 1971 (30).

service. But what alternative is there to these traditional monistic accounts? My central aim here is to develop an alternative in the form of pluralism about the good.

What precisely is pluralism about the good? It will help to contrast it with both monistic accounts of the good and with what is sometimes called the objective-list view. Pluralism holds that there are various goods, such as pleasure, knowledge, rationality, and virtue, and that there is nothing underlying or tying all of these together other than their intrinsic goodness. Monistic accounts, by contrast, claim that there is one single intrinsic good. Hedonism is a paradigmatic case of monism about the good. On most interpretations, hedonism claims that there is one type of good – pleasure or freedom from pain. Any other putative goods are valuable only insofar as they are useful in promoting, producing, or otherwise bringing about pleasure and freedom from pain. Another instance of value monism is perfectionism, according to which the good consists in the development or perfection of some aspect of human nature. For example, one contemporary perfectionist, Thomas Hurka, argues that knowledge, practical rationality, and certain kinds of physical or athletic achievements constitute the development or perfection of human nature. These are the components or constituents of a good life, and what unifies these apparently disparate goods, on Hurka's view, is their role in helping us to realize our true essence or nature as human beings. A rival version of perfectionism, associated with T.H. Green and F.H. Bradley, understands the good as the development of the self. On this monistic account, it is the concept of self-realization which is meant to unify our various judgments about the good. For example, Green suggests that we achieve self-realization through the development and exercise of those capacities which are essential to us as rational agents. On this view, all goods owe their favored status to their role in helping us to develop our true selves.

Pluralism denies that such monistic accounts can tell a plausible story about the human good. A recurring problem for monism in its various guises is its incompleteness. Hedonism asserts that only certain types of mental or experiential states can contribute to a good life. However, it is easy to imagine a life of intensely pleasurable experiential states which is nevertheless lacking either in actual achievements or in the exercise of one's rational capacities, and it seems undeniable that these missing elements detract from the value of such a life. There are also questions about hedonism's status as a genuine form of monism about the good. Sometimes hedonism is interpreted as holding that pleasure constitutes a sort of supervalue to which all other putative goods are to be reduced. However, as we shall see, it is not at all clear that pleasure is a unitary concept; indeed, the range of types of experiences which we can sensibly label pleasurable is quite diverse. Moreover, although hedonists sometimes equate pleasure and freedom from pain, or at least lump the two together, these two experiential states seem quite disparate. Thus, even hedonists seem ultimately committed to a form of value pluralism, albeit one where the good is restricted to the having of certain experiential states.

Perfectionists also suffer from an incompleteness problem; however for them the problem is one of making room for what we might call subjective goods. Hurka, for instance, admits that pleasure and freedom from pain have no role as intrinsic goods in his perfectionist account. Any value such states have is purely instrumental. Since it is easy to imagine fairly high levels of Hurkian perfectionist achievement with little in the way of pleasure and contentment, and a good deal in the way of pain and anxiety, this perfectionist account has some decidedly implausible implications. Moreover, as we shall see, Hurka's view in fact has trouble unifying his various perfectionist goods. He appeals to the notion of a human essence,

claiming that theoretical rationality, practical rationality, and what he calls physical perfection are all part of human nature. However, it seems very doubtful that there is any such thing as a human essence, and even if there is, it is extremely unlikely that the states favored by Hurka are part of that essence. Thus, just as the hedonist has trouble giving a genuinely monistic account due to the apparent disparateness of pleasures, Hurka's appeal to human nature and the human essence is incapable of underwriting his conclusions about the human good as well. Even Hurka's relatively short list of goods, therefore, seems to commit him to a form of pluralism about the good, and this simply becomes more obvious if we add subjective goods such as pleasure and freedom from pain to Hurka's list.

The self-realization version of perfectionism avoids some of the problems associated with Hurka's view. Whereas Hurka makes claims about the essence of human beings understood in biological terms, Bradley and Green are concerned not with the development of human nature, but instead with the development of the self, which they understand in decidedly non-biological terms. Bradley sees the self as being socially constituted in an important way. The beliefs, conventions, feelings of obligation, and ways of thinking of one's community are part of one's self, such that an attack on one's community is in fact an attack on oneself. Thus, for Bradley an important aspect of self-realization is the fulfillment of those duties associated with one's position in society; thus, on Bradley's view I am to find self-realization largely in the fulfillment of *My Station and its Duties*. Bradley is interesting for my purposes as a case study in how monistic accounts of the good very naturally gravitate toward a more plausible pluralism. Bradley recognizes that one's good cannot consist simply in the fulfillment of the duties associated with one's station in society. Thus, he recognizes other goods, such as the pursuit of scientific knowledge, artistic

achievement, something he calls ideal morality, and even pleasure. Although Bradley clings to the concept of self-realization as a unifying force in his account, we shall see that this attempt is in vain. Bradley offers no plausible understanding of the concept of self-realization which can unify the disparate goods which he is driven to recognize.

Green avoids some of this by being somewhat less explicit than Bradley in specifying the constituents of the good, and by eschewing any talk of a socially constituted self. For Green, the notion of the self is intimately connected to the idea of rational agency. The self is that part of us with the capacity to distance ourselves from our present desires and our doxastic impulses, and to determine what we ought to do or believe in light of our best reasons (either epistemic or practical). The development of the self is a matter of developing those capacities associated with epistemic and practical rational agency. Thus, for Green self-realization or the good is constituted by the development and exercise of those capacities essential to us as rational agents. This identification of the good with the development and exercise of our rational agency capacity is superior to Bradley's view insofar as it avoids certain implausible claims about the social nature of the self. Also, because it does not explicitly recognize other goods such as pleasure, artistic achievement, and so forth, Green's view seems better able to lay claim to the title of monistic account. Moreover, the notion of the proper exercise of our capacity for practical rational deliberation is just vague enough to promise a potentially longer and more diverse list of goods than other monistic accounts are able legitimately to deliver. However, as we shall see, this apparent strength of the Green view is actually one of its main weaknesses. There are several distinct ways of understanding the notion of the exercise of our capacity for practical rational deliberation, and none of them is ultimately defensible.

Although neither hedonism nor perfectionism offers an adequate account of the human good, there is something positive to be gained from a consideration of these views, for each view captures an important aspect or component of the human good. One of the reasons that so many have found hedonism to be a plausible account of the human good is that it seems genuinely undeniable that certain experiential states are intrinsically desirable, while others are intrinsically undesirable. Any plausible theory of the good must account for this. However, the notion that a good life can consist merely in the having or avoidance of such experiential states is an idea which does not withstand scrutiny, and the various forms of perfectionism recognize this. Any adequate account of the human good will have to recognize that, while human beings share a good deal in common with other animals, we are different from them in significant ways. Most importantly, we have a far greater capacity for rational deliberation. Although rationality is not *essential* to human nature, our rationality is nevertheless an important feature of normal human beings, and any adequate account will have to include the development and exercise of our capacities for rational deliberation as an important aspect of the human good.

An important truth at the heart of perfectionism is the notion that the good of an individual depends, in some sense, on the sort of being that it is. I think that hedonism and various versions of perfectionism suffer from an inadequate appreciation of the complexity of human nature. Hedonism takes a theory of the good which seems quite appropriate for non-rational animals and attempts to apply it to human beings. Thus, the oft-repeated criticism that hedonism is a doctrine fit more for swine than humans. Perfectionists, by contrast, tend to err in the opposite direction, proposing accounts of the good which seem more appropriate to disembodied rational agents. Although rational agency is central to our understanding of

our own nature, we should not lose sight of the things we have in common with the rest of the animal kingdom. Thus, the goods associated with rational agency, such as knowledge and practical rationality, must be supplemented by and weighed against the goods and evils associated with the body, such as pleasure and pain.

It is also important to distinguish value pluralism from a related view known as the objective-list account of the good. Some of what attracts certain people to monistic theories of the good is the promise of such views to unify and explain all of our ethical judgments. Monistic accounts seem better able to exemplify familiar theoretical virtues such as unity, simplicity, and explanatory coherence. Hedonism purports to explain all of our judgments about the good in terms of pleasure and the avoidance of pain – e.g., practicing the violin is worthwhile because of the pleasure promised by mastery of the instrument. Similarly, perfectionism explains all of our individual judgments about the good in terms of certain actions or activities exemplifying human nature or rational agency. At the opposite end of the theoretical-virtues spectrum is the objective-list account, according to which there are certain things which are objectively good or valuable. This list could include any number of things, such as knowledge, social recognition, moral virtue, artistic appreciation, athletic achievement, autonomy, pleasure, desire satisfaction, and so on. In some ways, this view resembles my own. I argue that there is a list of goods, and that there is no underlying supervalue or strong principle tying them all together. My list includes various kinds of pleasures and enjoyments, freedom from various kinds of pains and sufferings, knowledge, the exercise of our capacity for practical deliberation, and proper emotional responsiveness.

I take it that what separates my pluralistic account from the objective-list view is the latter's tendency to completely eschew any attempt at explaining or justifying the items on the

list. We are simply presented with a list of goods. To the extent that the objective-list account takes such a position, it is distinguishable from value pluralism. I argue that there are things we can say about why certain items get on the list of human goods. Moreover, although I am deeply suspicious of the sorts of unifying principles proposed by traditional versions of value monism, I nevertheless think that the kernel of truth in perfectionism—that one's good depends in some way on one's nature—does offer some degree of unity and explanatory coherence. Human beings have a complex nature. We are best thought of as embodied rational agents, and any plausible and adequate account of the human good will have to be correspondingly complex; it will have to take into consideration the sensual, rational, conative, and emotive aspects of human nature. The kind of theoretical unity and simplicity sought by traditional forms of value monism is simply not a realistic goal given our complex nature.

A natural response to the inadequacy of monistic accounts of the good is to propose a desire-based theory. One might very well agree that there are many goods, and that there is nothing they have in common other than their goodness, but insist on a version of the desire-satisfaction account of the good to explain this phenomenon. According to such an account, what's good for someone is the satisfaction of her desires—either the desires she actually has or, according to more sophisticated versions of the account, the desires she would have under suitably idealized conditions. Although a desire-satisfaction account would not suffer from the incompleteness problem shared by various types of monism, the view is nevertheless untenable. One problem with desire-based accounts is their distorted conception of the relation between our desires and what's good. Our experience of value is such that it is typically much more natural to describe ourselves as desiring something because it is good, or

because we believe it to be good in some way. The desire-satisfaction view instead entails that something is good because it is desired, or would be desired under certain conditions. I shall argue that, once we allow for the fact that the satisfaction of our desires typically results in pleasure and other objective goods, while the frustration of desire typically results in feelings of frustration and other objective evils, the idea that our desires can *make* something good or bad for us is extremely implausible. A second problem for the desire-satisfaction view is, regardless of how we specify the account, it nevertheless has implications so counterintuitive as to render the position indefensible. The third problem for the view is that, particularly in its most sophisticated version, it lacks the sort of concrete guidance one expects from a theory of the good. Peter Railton, for instance, defines the good as that which a fully informed and completely rational version of oneself would recommend to one's less informed and imperfectly rational self. The problem is that it is hard to know what the fully informed and rational self would recommend unless we first give some serious thought to the issue of what is good. Thus, Railton's view still leaves us without a substantive account of the human good.

At this point a few words about my methodology would perhaps be helpful. I like to think of myself as operating in the tradition of Aristotle and Henry Sidgwick, both of whom saw the consideration of the ethical views of others as an important starting point in developing their own accounts of virtue, the good, and the right. Accordingly, in this dissertation I spend a significant amount of time systematically considering what I take to be the most important candidates for a theory of the good. Also in the spirit of Aristotle and Sidgwick, I try to extract what is of value from the various philosophical accounts which turn out to be, for one reason or another, untenable. Moreover, like Aristotle I do not limit myself

to a consideration of philosophical sources only; I also pay some attention to more popular sources, since I agree with Aristotle that it is important to consider not only the beliefs of the wise, but also those of the many.³ Thus, I will from time to time make use of examples taken from popular culture. Sometimes this is because such examples allow for an especially clear illustration of a particular problem or view which I want to discuss. But it will also demonstrate that, even in that part of our culture which one would expect to be most receptive to subjective accounts of the human good, there is at the very least considerable ambivalence about the supreme or overriding value of pleasurable or otherwise enjoyable experiential states. Moreover, although the question of what constitutes a good or worthwhile life is by no means an easy philosophical problem, it is nevertheless one which, in contrast to some of the more esoteric problems of philosophy, presents itself in an often inescapable way to even the most ordinary of people. Thus, although popular culture might not be a particularly fertile source for ideas about the metaphysical status of time, it can be useful in developing an account of the human good.

I also somewhat self-consciously see myself as operating within the tradition of reflective equilibrium – a phrase made famous by Rawls in *A Theory of Justice*. The model of reflective equilibrium has its roots in the approach of philosophers such as Aristotle and Sidgwick, and since Rawls' explicit statement of it more than 30 years ago it has become the dominant methodological paradigm for philosophical ethics, particularly in the English-speaking world. Here is a clear description of the approach by a pair of prominent biomedical ethicists who are sympathetic with the reflective-equilibrium model:

[Rawls] views justification as a reflective testing of our

³ See, e.g., *Nicomachean Ethics* 1098b, 25-30.

moral beliefs, moral principles, theoretical postulates, and other relevant moral beliefs in order to make them as coherent as possible. Method in ethics properly begins with our considered judgments, the moral convictions in which we have the highest confidence and believe to have the lowest level of bias. The term *considered judgments* refers to judgments in which our moral capacities are most likely to be displayed without distortion. Examples are judgments about the wrongness of racial discrimination, religious intolerance, and political repression. These considered judgments occur at all levels of generality in moral thinking, from those about particular situations and institutions through broad standards and first principles to formal and abstract conditions on moral conceptions.⁴ Whenever some feature in a moral theory that we hold conflicts with one or more of our considered judgments, we must modify one or the other in order to achieve equilibrium.

Even the considered judgments that we accept provisionally as fixed points are, Rawls argues, liable to revision. The goal of reflective equilibrium is to match, prune, and adjust considered judgments in order to render them coherent with the premises of our most general moral commitments. We start with paradigm judgments of moral rightness and wrongness, and then construct a more general and more specific account that is consistent with these paradigm judgments, rendering them as coherent as possible. We then test the resultant action-guides to see if they yield incoherent results. If so, we readjust these guides or give them up and then renew the process.⁵

Each of us has a set of ethical beliefs or commitments at varying levels of generality. We are more strongly committed to some of these than to others. Not surprisingly, we often find that some of these ethical judgments conflict with one another. Often the conflict is between a general ethical principle and some more specific or particular ethical judgment. For instance,

⁴ This quote is taken from Rawls' *The Independence of Moral Theory*.

⁵ Tom Beauchamp and James Childress, *The Principles of Biomedical Ethics*, 5th edition p. 398. The roots of Rawls' notion of reflective equilibrium can be found in his 1951 essay *Outline of a Decision Procedure for Ethics*, though apparently he did not use the expression *reflective equilibrium* until *A Theory of Justice* (See 20f, 48-51, 120, 432, 434, 579-80).

suppose one is committed to the general principle that lying is always morally impermissible, but when faced with a situation where telling the truth would result in the death of an innocent person, judges that lying is permissible in this case. Clearly, there is tension between the general belief about the moral status of lying and the particular judgment about the permissibility of lying in this situation. An obvious solution to this problem is to give up or modify the general principle, so that perhaps one now holds the alternative general principle, One should always tell the truth unless someone's life is at stake. This new general principle may or may not be adequate—perhaps there are other cases short of lives being at stake where we will say that it is permissible to lie. If so, we will be forced to give up this particular judgment or further modify our general principle. The process of reflective equilibrium requires that we go back and forth between general principles and particular judgments, altering and modifying them, until we reach a state where our various judgments are in equilibrium. This is what we are doing when we attempt to develop an adequate account of some aspect of our ethical life, and it is the model I make use of in developing an account of the human good.

Implicit in the reflective-equilibrium model is a push toward greater explanatory coherence and systematicity in our ethical thinking. These properties are important since they tend to give an account more power to criticize or otherwise inform our particular ethical judgments. Rawls proposed the reflective-equilibrium approach as the methodological foundation for an alternative to both utilitarianism and what he called intuitionism. Rawls was mainly interested in developing an account of justice, and the sort of intuitionism he had in mind was what one might call moral pluralism, according to which there are various competing fundamental moral principles, which can and often do conflict with one another,

and none of which can be reduced to some supreme moral principle such as the principle of utility. Any reader familiar with Rawls' discussion of intuitionism in *A Theory of Justice* and his desire to avoid this sort of pluralism might think that there is some obvious tension in my acceptance of the reflective-equilibrium methodology along with pluralism about the good. However, we need to keep in mind that, notwithstanding the model's tendency to push us in the direction of greater systematicity, unity, simplicity, and explanatory coherence, it is nevertheless an open question just how much of these various theoretical virtues a plausible account of the good can achieve. For any theory of the good will have to make room for certain considered judgments about the good, and while no adequate account can tolerate tension or contradiction, there is a significant amount of fuzziness and indeterminateness which one might find in an adequate account of the human good. Indeed, I shall argue that pluralism is the view which best survives the process of reflective equilibrium.

The form of pluralism I propose recognizes no strong ordering principles among the various goods. For example, I see no reason to think that any particular good such as pleasure or knowledge has anything like lexical priority over other goods. Moreover, I reject the claim that the various goods are conditional in some way – e.g., that pleasure is good only if it is morally innocent or that a perfectionist good such as knowledge or artistic achievement is valuable only if it is endorsed or desired. Also, less controversially, I am skeptical about our ability to quantify in a robust, publically defensible way the value of various goods. Although I do think we make judgments about the relative value of particular goods or ways of life, such judgments are of a highly intuitive and imprecise nature. Judgments about what would be best for someone all things considered are messy, and it's not at all surprising that so many philosophers have gravitated toward tidier monistic accounts. However, while such monistic

theories have the advantage of simplifying our judgments about the good, this comes at an extremely high cost. Any monistic account forces us to ignore important aspects of the good. In effect, we are asked to pretend that the good life consists in the pursuit of one particular aspect of the good. We are expected to arbitrarily pick one particular type of good and live as if other, competing goods are simply not important or worth pursuing. Monistic accounts such as hedonism and perfectionism are really nothing more than intellectual fetishes. They ignore the complexity of human nature and the human good, and are recipes for impoverished lives. Stay away from them.

This dissertation is organized in the following way. Chapter 2 deals with what I call subjective theories of the good, which include hedonism and the desire-satisfaction view. Chapter 3 considers an important contemporary version of perfectionism proposed by Thomas Hurka. In Chapter 4, I examine so-called self-realization accounts of perfectionism. These views are most closely associated with Green and Bradley and offer an interesting contrast to Hurka's version of perfectionism. In each of these chapters I explain why the account under consideration is untenable, and reveal what I think is correct about each view. It is in chapter 5 that I offer my own pluralistic account of the good; I identify some intrinsic goods which any adequate account must recognize, and articulate my reasons for rejecting certain constraints on our judgments about the good. In the end, we'll find that there are many things which can go into making a good life, but that these things can and often do conflict with one another. Not surprisingly, planning for a good life consists in making sound but nevertheless intuitive judgments about the relative value of the competing goods, and being informed and realistic about one's own capacities and resources for achieving and experiencing various goods. To live a good life, one must know what's good and one must know oneself.

Chapter 2 Subjective Theories of the Good

In this chapter, I argue against subjective theories of the good. There are two distinct versions of subjectivism: hedonism and the desire-satisfaction view. According to hedonism, the only intrinsic good is pleasure and freedom from pain; any other putative goods such as knowledge and virtue are of mere instrumental value. The desire-satisfaction view, by contrast, holds that the good consists in the satisfaction of one's desires—either one's actual desires or the desires one would have under certain idealized conditions. While the desire-satisfaction view is the more popular of the two versions of subjectivism, at least among contemporary philosophers, hedonism is the older of the two doctrines and has had more than its fair share of adherents throughout the history of philosophy. Moreover, hedonism seems to retain a certain allure or *prima facie* plausibility among non-philosophers; thus, novice philosophy students tend to gravitate toward hedonism when they are prodded to offer a theory of the good. Although the popularity of these subjective accounts is understandable, they are nevertheless untenable. I begin by considering hedonism, and then turn to the desire-satisfaction view.

There are some important interpretive issues regarding hedonism which must be addressed if we are to properly evaluate the view. The first question is whether hedonism is best thought of as a form of monism or pluralism about the good. Monists hold that there is one type of thing which is intrinsically valuable, while pluralists contend that there is a diverse set of goods having nothing in common other than their intrinsic goodness. A second question is whether hedonism is best thought of as a form of subjectivism or objectivism about the good. Although hedonism has traditionally been classified as a form of

subjectivism, there are interpretations of the view according to which it is an objective account of the good. Moreover, there are at least two distinct ways in which hedonism could be interpreted as a form of subjectivism. Different interpretations of hedonism will rely on different arguments and will be vulnerable to different objections, so in considering hedonism it will be important to be clear about how we are interpreting the view. My strategy will be to consider what I take to be the strongest and most influential arguments in favor of hedonism, addressing the relevant interpretive issues as I go along.

One might be attracted to hedonism because of its apparent monism about the good. Many are attracted to maximizing conceptions of practical reason, according to which self-interest or morality is a matter of bringing about the most good either for oneself or for some defined class of individuals. It is natural for such views to be paired with conceptions of the good according to which there is one particular type of thing possessing intrinsic value. Monistic accounts of the good promise a straightforward method for determining what would be best from either a moral or prudential point of view: it is simply a matter of determining what would bring about the greatest amount of the one intrinsic good. Since pluralistic accounts of the good require the weighing and balancing of quite disparate goods such as pleasure and knowledge, monistic accounts are often seen as offering more in the way of theoretical simplicity, unity, and determinacy. Where a pluralist requires a judgment about the relative value of such disparate goods as pleasure, knowledge, artistic excellence, and so forth, the monistic hedonist only has to ask what would bring about the most pleasure.

Although there are interpretations of hedonism which can accurately be described as monistic, such versions are implausible both as comprehensive accounts of pleasure and as theories of the good. The sort of hedonism favored by the Epicureans seems to be a monistic

account. Epicurus and his followers famously held pleasure and the absence of pain to be the final and ultimate good, that to which every choice or action is a means, and which itself is not a means to anything.⁶ Anything else which is good – e.g., wealth, power, social prestige, virtue, philosophical reflection, etc. – is so due to its being a means to pleasure or freedom from pain. The Epicureans typically held the absence of pain to be the greatest pleasure.⁷ For the Epicurean, the best life is one of freedom from pain in the body and from disturbance in the soul.⁸ Moreover, Epicureans typically held that disturbances of the soul were worse than bodily pains, and that the greatest pleasures were those of the mind.⁹ We should note here that this form of hedonism seems to constitute an objective theory of the good. The Epicureans do not say that the good is whatever one desires; instead, they identify the good with a certain kind of mental state – one which could fairly be referred to as tranquility.¹⁰ If one ultimately desires some state other than tranquility – e.g., the exciting but harrowing experiences of a soldier – then one is living a life inferior to that of the Epicurean sage. Even if the soldier takes his experiences to be good for him, and prefers his own life to one of quiet contemplation, he is nevertheless wrong. He, like anyone else, would be better off living a life free of the sort of mental disturbances he has come to value as a soldier.

It is not obvious, however, that tranquility and freedom from bodily pain and mental

⁶ See for example Cicero's *On Ends*, excerpted in Long and Sedley 21A(1).

⁷ See Cicero's *On Ends*, in Long and Sedley 21A(6).

⁸ Epicurus, *Letter to Menoeceus*, in Long and Sedley 21B(5).

⁹ See Diogenes Laertius in Long and Sedley 21R(2), and Cicero's *On Ends* in Long and Sedley 21U(2).

¹⁰ Long and Sedley translate the Greek word *ataraxia* as tranquility (see p. 124). This tranquility, or freedom from mental disturbance, is taken by Long and Sedley to be the supreme hallmark of Epicurean happiness (p. 124).

disturbances constitute a single type of state. For entities lacking consciousness clearly are free of such states, yet we would not say that they possess tranquility. However, one can amend the view by requiring that the individual which is free of pain and agitation be a conscious being. In that case, it would seem considerably more plausible to identify tranquility with freedom from bodily pain and mental disturbances. However, now we are left with an implausibly narrow conception of pleasure. Although tranquility is a type of pleasure, there are others. We think of ourselves as taking pleasure in a wide range of things; however, it does not seem that these various experiences have anything in common, other than our positive attitude toward them or their intrinsic desirability. Consider the pleasure that one takes in the following activities or events: leading your team to the championship; eating a banana split; witnessing the birth of your first child; successfully working through an especially difficult mathematical proof; building and maintaining a close friendship; riding a roller coaster; listening to an interesting academic lecture; having a warm bath after being caught out in an ice storm. These are all things that one normally expects to be pleasurable; however, there does not seem to be any phenomenological feature that they all have in common which would make them all instances of pleasure. Indeed, the list seems strikingly diverse; e.g., the pleasure one finds in riding a roller coaster seems all together different from the pleasure derived from dozing in a hammock on a warm summer afternoon.

The same can be said about the diverse set of experiences which we think of as painful. Although the physical pain from a toothache, fear of death, anxiety about an impending speaking engagement, anger about some great injustice, humiliation by a bully in front of a group of friends, extreme hunger, disappointment about flunking out of medical school, and the ache from a stubbed toe are all disutilities, and in some sense painful, it is

hard to see what phenomenological feature they all have in common. Although we often talk of mental states such as fear, anxiety, and depression in terms of psychological pain, on reflection it would seem that these often have little or nothing in common with the sorts of pain caused by such things as extreme hunger or a gunshot to the kneecap, other than our negative attitudes toward them or their intrinsic undesirability. Moreover, these differences are not merely differences in *quantity* of pain; it seems instead that the pain from a gunshot wound is *qualitatively* different from a fear of public speaking.¹¹

Although the notion of pleasure as a monistic state has some prima facie plausibility, it does not withstand critical scrutiny. Pleasures and pains are quite diverse, and even if we think that the good consists in pleasure and freedom from pain, the diverse nature of these states will still leave us with a form of value pluralism. Thus, we will be forced to make judgments about the relative value of diverse goods, and therefore hedonism cannot be justified by appeal to the theoretical virtues of monistic accounts of the good. Moreover, any form of hedonism which stipulates that one particular type of pleasure such as tranquility is the sole good will offer an impoverished, arbitrary, and narrow conception of the human good and thus ought to be rejected.

In order to evaluate our other arguments for hedonism, it will help to be clear about the various ways in which hedonism might be thought to be a subjectivist account of the good. Any form of hedonism will make essential reference to mental states, and since mental states are subjective in a straightforward sense, there is straightforward sense in which hedonism is necessarily subjectivist. However, there is a more interesting way in which hedonism might

¹¹ A number of philosophers, including many utilitarians, have noted this problem for hedonism. See for example Sidgwick 1907 (127), Griffin (8), and Brandt (258).

or might not be thought of as subjectivist. Both Jennifer Whiting and Richard Arneson have given reasons for classifying hedonism as a form of objectivism. Whiting includes all views (such as hedonism) which identify the good with some subjective psychological state (such as pleasure) as objective in so far as they claim that pleasure (or some other subjective psychological state) is good for us independently of our belief that it is so that is, in so far as they allow that someone may...mistakenly believe that pleasure is not in fact good for him, and so, mistakenly avoid pleasure (Whiting: 1988: 46 n3). Similarly, Arneson criticizes Derek Parfit's division of theories of the good into the categories of hedonism, desire satisfaction, and objective-list theories, suggesting that hedonism, notwithstanding its traditional classification as a form of subjectivism, can be construed as a version of the objective-list theory if we assume that pleasure is the only item on the list of objective goods (Arneson 1999: 115). Arneson's solution is to define subjective theories as those that claim that what is good for each person is entirely determined by that very person's evaluative perspective (1999: 116). Put another way, objective theories accept, and subjective theories deny, the claim that what is good for a person can be correct or incorrect and that the correctness of a claim about a person's good is determined independently of that person's volitions, attitudes, and opinions (Arneson 1999: 115).

With Epicureanism, we have already seen one version of objective hedonism. Also, Henry Sidgwick seems to have held an objective version of hedonism, or at least an objective version of what we might call a *quality-of-experience* view, according to which the good consists solely in certain kinds of intrinsically desirable mental states, which may or may not be classified as pleasures. Although Sidgwick rejects the idea that there is some phenomenological feature which all purported pleasures have in common, he nevertheless

thinks that certain mental states are intrinsically desirable:

Shall we then say that there is a measurable quality of feeling expressed by the word pleasure, which is independent of its relation to volition, and strictly undefinable from its simplicity? like the quality of feeling expressed by sweet, of which also we are conscious in varying degrees of intensity. This seems to be the view of some writers: but for my own part, when I reflect on the notion of pleasure, using the term in the comprehensive sense which I have adopted, to include the most refined and subtle intellectual and emotional gratifications, no less than the coarser and more definite sensual enjoyments, the only common quality that I can find in the feelings so designated seems to be that relation to desire and volition expressed by the general term desirable, in the sense previously explained. I propose therefore to define Pleasure when we are considering its strict value for purposes of quantitative comparison as a feeling which, when experienced by intelligent beings, is at least implicitly apprehended as desirable or in cases of comparison preferable (Sidgwick 1907: 127).

Sidgwick goes on to say that what is good for an agent ultimately is intrinsically desirable feelings or consciousness (397-98). He also says that the desirability of each feeling is directly cognizable by the subject of the experience (398). Whereas the Epicureans identified one particular type of mental state which they claimed was intrinsically valuable (tranquility), Sidgwick recognizes that there are many such intrinsically desirable states, and the only thing they have in common is their intrinsic value or desirability, which is directly cognizable or implicitly apprehended by the agent.

Clearly, Sidgwick thinks that there is an important relation holding between the property desirableness and desire and volition. Sidgwick defines the good as that which is desirable, and identifies desirability with what is reasonably desired (xxviii). Sidgwick says that a man's future good on the whole is what he would now desire and seek on the whole if all the consequences of all the different lines of conduct open to him were accurately

foreseen and adequately realised in imagination at the present point of time. (111-12). One might be tempted to take these statements as evidence of a subjectivist version of hedonism in Sidgwick's thinking about the good; however, I think that would be a mistake. Sidgwick accepts the claim that whatever is desirable would in fact be desired by a fully rational and fully informed agent. However, this does not mean that certain states are valuable or desirable *because* they would be desired under certain idealized conditions. Instead, certain states are by their very nature intrinsically desirable, and their intrinsic desirability is directly cognizable, on Sidgwick's view.

Sidgwick's rejection of dispositional accounts of desirableness is most obvious when we consider his claims about the authoritativeness of our judgments about the desirability of our own experiential states. According to Sidgwick, the desirability of each feeling is only directly cognizable by the sentient individual at the time of feeling it, and...therefore this particular judgment of the sentient individual must be taken as final on the question how far each element of feeling has the quality of Ultimate Good (398). He also says that no one is in a position to controvert the preference of the sentient individual, so far as the quality of the present feeling alone is concerned (128). It is difficult to square what Sidgwick says about the certitude of our judgments about the desirability of our own mental states with his equating of a man's good with what he would now desire and seek on the whole if all the consequences of all the different lines of conduct open to him were accurately foreseen and adequately realised in imagination at the present point of time. For it seems that I could be quite mistaken about what I would desire under such idealized conditions. However, there is another way of reading Sidgwick's claim. I think we should take him as claiming that, while it is always directly discernable whether a given feeling or state of consciousness is desirable,

we often cannot accurately predict what sorts of mental states will accompany alternative courses of action. Thus, my future good is identical with what I would desire under conditions of full information and rationality; however, knowledge of all the consequences of all the different lines of conduct open to me along with an adequate imaginative representation of these consequences does not tell me anything about the *desirability* of those accompanying mental states. Desirability of mental states is directly cognizable by the subject of experience at the time of the experience. Having full information and rationality allows me to predict which mental states will ultimately accompany which courses of action available to me, and thus allows me to determine what course of action would best contribute to my good. An immediate problem for Sidgwick's account is how he is to justify his claim that desirable consciousness is the only thing of intrinsic value. He claims that reflection on the question reveals that this is the case (113, 395-97). However, this seems doubtful, as we shall soon see.¹²

There are at least two distinct ways in which a hedonistic account of the good could be a form of subjectivism. First, one might accept a non-Sidgwickian, but nevertheless objective, conception of pleasure, according to which pleasure consists in some unitary mental state or perhaps in any experiential state accompanied by what some have called a hedonic tone.¹³ On this view, what makes the possession of such mental states good for the agent is the agent's desire for them. In other words, what is good for someone is the satisfaction of her desires; and we end up with hedonism by combining this claim about the good with the empirical or psychological claim that agents as a matter of fact intrinsically

¹² This will be brought out most clearly in our discussion of Nozick's experience machine.

¹³ See for example Broad (229).

desire only this certain sort of mental state. This is one way of taking arguments which attempt to move from the alleged truth of psychological hedonism to normative or ethical hedonism.¹⁴ There are several problems with such a view. First, as we have already seen in our discussion of monism, the claim that pleasure consists in some unitary mental state is simply implausible. Second, the claim that pleasure, however one defines it, is the only thing we intrinsically desire seems empirically false. Finally, and perhaps most importantly, this version of hedonism collapses into the desire-satisfaction view, and thus ceases to be an interesting account of the good in its own right. Any evaluation of this form of hedonism will depend on our evaluation of the more general claim that one's good consists in the satisfaction of one's actual or hypothetical desires.

A second way in which one can be a subjectivist hedonist is by accepting a subjectivist or attitudinal account of the nature of pleasure. According to such a view, pleasures are those mental states which we have some sort of pro-attitude toward, such as enjoying, liking, or desiring to continue. This subjective definition of pleasure is detachable from a desire-satisfaction account of the good; one who accepts the former does not have to be committed to the latter. However, while a view which combines the rejection of a desire-based conception of the good with an attitudinal account of the nature of pleasure is a logically possible position, it is difficult to see the motivation for it. Especially if we make the plausible assumption that pleasure is intrinsically good, it seems to follow that it is our attitudes which make pleasure a basic or intrinsic good. If our attitudes can confer basic or intrinsic value onto mental states, it is hard to see why we should draw the line there and not

¹⁴ See Sumner (83-87) for a discussion of the distinction between psychological and ethical hedonism. See Chapter IV of Mill's *Utilitarianism* for an example of an appeal to psychological hedonism in support of ethical hedonism.

also say that our pro- attitudes can make other things good for us as well. But to accept this conclusion would be to commit us to some version of the desire-satisfaction view; and again, hedonism would cease to be an interesting or distinctive view in its own right. Thus, our evaluation of this version of hedonism would depend on the plausibility of the desire-satisfaction view.

Richard Brandt represents a version of hedonism which combines the desire-based conception of the good with an attitudinal conception of pleasure. Unlike some hedonists, Brandt draws a clear distinction between hedonism (what he sometimes calls the happiness theory) and the desire-satisfaction view (246). The happiness theory holds that one's good consists in a certain kind of mental state, which we refer to as happiness; on this view, the rationally benevolent person acts in such a way as to maximize happiness for people or, more generally, for sentient beings. The desire-satisfaction view, by contrast, holds that what is good for a person is the satisfaction of her desires, as opposed to happiness; on this view, the rationally benevolent person works to see that people or, more generally, sentient beings, get what they want as opposed to happiness or enjoyment *per se*. Brandt gives two reasons for holding the happiness view:

First, that what we care about securing for other persons (e.g., our own children) is their happiness; and we seem to care about getting them what they want (or what they would want if they knew more, etc.) only to the extent we think that so doing will bring them happiness or avoid distress and depression. Second, the psychological theory of benevolence leads to the conclusion that what we are sympathetically motivated to secure for others is happiness and freedom from distress (although we may want desire-satisfaction because we believe it a means to these) (248-249).

Here, Brandt seems to suggest that happiness has some status and value for us independently of our desires or volitions; thus, his view could be mistaken for a form of objectivism about

the good. However, we should note that Brandt understands happiness or pleasure in attitudinal terms.¹⁵ Moreover, he thinks that happiness is good for us *because* it is what we ultimately desire. Brandt wants to show how we can rationally evaluate or criticize desires or preferences (110-11). On Brandt's view a desire is rational if and only if it would survive what he calls "cognitive psychotherapy" — i.e., if and only if the desire would be present in some persons if relevant available information registered fully, that is, if the persons repeatedly represented to themselves, in an ideally vivid way, and at an appropriate time, the available information which is relevant in the sense that it would make a difference to desires and aversions if they thought of it (111). Thus, although Brandt opts for a version of hedonism which holds that the ultimate good is the same for everyone (i.e., happiness), his view is at its roots a form of subjectivism. For happiness derives its value for us from the fact that a desire for it would never extinguish under cognitive psychotherapy; happiness (or pleasure or enjoyment) is what we ultimately desire.

I will address that part of Brandt's argument dealing with the desire-satisfaction view in a moment, when I turn to a systematic consideration of various versions of the desire-satisfaction account. For now, I would like to consider Brandt's claim that happiness or pleasure is what we are ultimately concerned to secure for those toward whom we are sympathetically disposed, such as friends or children. Although Brandt is quite correct to suggest that the promotion of pleasure, happiness and other positive experiential states is an important aspect of benevolence, and it is undeniable that much self-interested behavior is directed at the gaining and avoidance of certain kinds of experiential states, it is simply false

¹⁵See Brandt's discussion of the motivational theory of pleasure (38ff., 83, 148, 258).

that this is *all* we care about, either for ourselves or for our loved ones. Robert Nozick's famous experience machine thought experiment¹⁶ shows that any such quality-of-experience view, regardless of how it is grounded, whether it is objective or subjective, monistic or pluralistic, cannot possibly capture our considered judgments about the good. Nozick asks us to imagine a futuristic machine to which we can be attached, which will simulate any experience one might imagine. Let's suppose I want to know what it would be like to strike out Al Simmons, Lou Gehrig, Babe Ruth, Jimmy Foxx, and Joe Cronin all in order, as Carl Hubbell did in the 1934 All-Star Game. All I would have to do is go to the experience machine, and I would have all the relevant experiences: the roar of the crowd, the feel of the rosin, the smell of the leather glove, the weight of my sweat-soaked wool uniform, the feel of the baseball in my hand as I prepare to throw my wicked Carl Hubbell-style screwball.

One could explore any number of experiences in this way—anything from the experience of a general such as Napoleon or Julius Caesar gobbling up territory to the experiences of a bored housewife in 1920's rural Minnesota. No doubt, many of us would find ourselves spending a good deal of time in the experience machine; after all, one of the reasons we read books such as Caesar's *Commentaries on the Gallic War* or Sinclair Lewis' *Main Street* is that they give us some idea of what life is like for other people (though this of course is not the *only* reason). However, the question is whether we would want to spend *all* of our time in Nozick's experience machine. Suppose the machine is sophisticated enough so that it can replicate all of the experiences one would associate with a reasonably good life. One might have the experience of going to college, having a successful marriage, raising a

¹⁶ Nozick (42-45).

healthy and happy family, enjoying success in one's career, and so on, while all along one is really just floating in a tank, with the experience machine providing the illusion of living this kind of life. Suppose further that one spent all of one's life attached to the experience machine. Would this be a good life? Is there any relevant difference between the illusion of a good or desirable life via the experience machine and actually living such a life.

Virtually all of us, I take it, would say that the life one leads through the experience machine is no life at all.¹⁷ We don't merely desire to have the *experience* of raising children, or caring for our parents in their dotage, or accomplishing something significant in our careers; we want actually to *do* these things. However, Sidgwick's view that nothing is desirable except desirable feelings would deny this, as would Brandt's claim that we seek the satisfaction of our loved ones' desires only to the extent we think that so doing will bring them happiness or avoid distress and depression.¹⁸ (398). If we accept Sidgwick's claim that the ultimate good is desirable consciousness, then we must conclude that the life led via the experience machine is just as good as the real thing, and this is a deeply counterintuitive result. Moreover, this objection is applicable to all versions of hedonism, regardless of how they conceive of pleasure or happiness, since all versions of hedonism hold that the good consists in a certain type (or types) of mental state, or a certain quality of experience. Thus, even if we accept the desire-satisfaction view, we don't get hedonism, since it is false that what we ultimately desire is a certain kind of consciousness. The experience-machine objection to hedonism is decisive.

¹⁷ It's not just philosophers who have this intuition about the experience machine. Two recent popular movies, *Matrix* and *Vanilla Sky*, both seem to embrace the idea that a real life in the real world is preferable to something like Nozick's experience machine.

¹⁸ Brandt (248); Sidgwick 1907 (398).

Is there anything in hedonism that is salvageable? We have seen that the human good cannot consist simply in pleasure or desirable consciousness. However, one might argue that although pleasure or desirable consciousness is not sufficient for a good life, it is nevertheless a necessary condition. There are two ways in which such a claim might be interpreted. The first is simply the claim that no life is happy (or good) unless it includes at least some pleasure or desirable consciousness; this is a strikingly modest claim, and not one that I would be willing to deny. However, one could take it as the much stronger claim that nothing can make our lives go better or worse unless it somehow affects the quality of our experience (Sumner 1996: 112). L.W. Sumner takes this to be kernel of truth in hedonism, claiming that any adequate theory of the good will have to connect our well-being in some way or other with the experience of the conditions of our lives (1996: 112).

However, this alleged truth of hedonism turns out to be false as well. It is not the case that for something to make my life go better or worse it somehow has to affect the quality of my experience. A couple of examples will help to show why. Let us say that one of the central goals in Abraham Lincoln's life was the long-term preservation of the Union. This was something for which Lincoln worked tirelessly, and for which he was prepared to make great sacrifices. The pursuit of this goal dominated a good part of Lincoln's adult life, particularly his professional life. When Lincoln was killed, the Confederacy had been defeated on the battlefield; however, it was not at all obvious that the Union could still be preserved in anything like its pre-war state. There was still the very real possibility of terrorist or guerilla campaigns against Union soldiers and pro-Union political figures in the South, and there was no guarantee that the policy of Reconstruction would be even remotely successful. We think of Lincoln as having led a good life not just a morally good or

virtuous life, but a life that was good *for him* largely because his central goal of preserving the Union was ultimately realized. But there were many things that could have gone very differently after Lincoln's death—e.g., the eventual dissolution of the Union or the long-term political subjugation of the South—which would have made his life less successful. Since these are all things that would have happened after Lincoln's death, none of them would have had any effect on Lincoln's quality of experience.

It's not just things that happen after our deaths that can affect our quality of life without our knowing about them. Suppose I have a wife and a circle of friends, all of whom I take to be devoted and loyal. Unbeknownst to me, my wife has engaged in a series of affairs and my friends actually despise me, only pretending to like me so that they can bilk me out of money through various schemes. Moreover, periodically they all get together and have anti-Jeff parties, at which they further conspire against me, and also laugh at embarrassing video tapes of me provided by my wife. Even if I am completely oblivious to all of these horrible things that go on behind my back, one would still conclude that these things make my life worse, and that my life would be better if they secretly had pro-Jeff parties, and whispered only kind words about me. It is not the case that for something to affect the quality of my life it has to affect my consciousness or quality of experience. Thus there is no truth to even Sumner's rather modest truth of hedonism.

What can one say in favor of the other major form of subjectivism about the good—the desire-satisfaction view? It is open to the subjectivist who recognizes the futility of hedonism to cling to desire satisfaction, holding that the good consists in whatever an agent desires, or would desire under certain circumstances. This version of subjectivism is not vulnerable to either of the two main objections I have raised against hedonism, and is in

general the preferred view among contemporary subjectivists. There are two versions of the desire-satisfaction view; however, neither is tenable.

One version of the desire-satisfaction view holds that the good consists in the satisfaction of whatever desires the individual agent actually has. This position is attractive for a number of reasons. First, this view identifies the good in such a way that we have some common currency in making interpersonal and intrapersonal comparisons of the good.¹⁹ Desire-satisfaction theorists typically define desires or preferences in quasi-behavioristic terms. For example, James Griffin argues that desires do not have to have felt intensities; they need not be linked exclusively with appetitive states (some are, but others are aims we adopt as a result of understanding and judgement); they need not have existed before fulfillment. Rather, desiring something is, in the right circumstances, going for it, or not avoiding or being indifferent to getting it (Griffin: 14). Brandt offers a similar definition of a desire: what it is for a person to desire some situation *O* occurrently, for its own sake, is for the person to be in such a central state that, were he at that time to think that performing a certain action would make the occurrence of *O* more likely, he would be more inclined to perform that act. In other words, if a person desires *O*, then, if he sees an available approach to *O*, he will be more inclined to move along that route (Brandt: 95-96). There is nothing mysterious about this conception of the good; one's preferences or desires are revealed by

¹⁹ Some versions of classical hedonism define pleasure in such a way that they are able to achieve the same result. As we have seen, the Epicureans identified *ataraxia* as the highest pleasure, and if one assumes that we have a reliable way of determining when someone is in a state of *ataraxia*, then there is no problem, in principle, with making interpersonal or intrapersonal comparisons. Similarly Jeremy Bentham held that the diverse range of pleasures and pains of which human beings are capable are reducible to ordinary physical pleasure and pain (see Sumner 109). I have argued that such analyses are implausible.

one's choice behavior, while we determine whether a given desire has been satisfied by looking to see whether the state of affairs that is the object of the desire has obtained.

A second advantage of the desire-satisfaction view is that, unlike hedonism, it does not locate the good solely in any mental state or quality of experience. Thus, it avoids Nozick's experience machine objection. What's good for me (according to the desire-satisfaction view) is whatever I desire, and since I often desire things other than certain kinds of feelings or consciousness, we can make sense of the notion that a life spent in the experience machine is inferior to one actually lived in the real world. Also, the desire-satisfaction view can account for our intuition that the quality of our lives can be affected by things that do not affect our experience. Lincoln desired that the Union be preserved; indeed, this was one of his most prominent desires, playing a central organizing role in his hierarchy of desires. If this desire is satisfied, then Lincoln's life, to that extent, goes better. The desire-satisfaction view makes no requirement that the agent with the desire be aware of its fulfillment.

The actual-desire view, with its emphasis on desires or preferences revealed by choice behavior, has been especially popular among economists and other social scientists.²⁰ Not only does it make for greater theoretical simplicity and explanatory unity, but it also makes it easier for economists to predict and explain consumer behavior, since desires or preferences understood in a behavioristic manner are publically observable in a way that states of consciousness or feelings are not. Moreover, the view seems to offer a defense against paternalism; for if one's good simply consists in the satisfaction of whatever desires one has, then it is hard to see how the government could justify interfering with individuals

²⁰ See Sumner 1996 (113-122) and Griffin (10).

making their own decisions about how to live their own lives. Thus, the actual-desire view also seems to offer support for something like Mill's harm principle.

However, notwithstanding the practical and theoretical convenience that the actual-desire view has for economists, and its potential support for anti-paternalism, the view has some very serious, and very obvious, problems. The most serious problem for the view is that it fails to capture the commonplace experience of our desires being mistaken. We have all had the experience of wanting something, getting it, and then finding that it was not what we wanted after all. This is not restricted to whims—e.g., one might desire a career in the military, with this desire playing a central motivating role in the life of the agent, but when the desire is satisfied, the agent finds that it was all a horrible mistake; life in the military is not a good life. Similarly, we often find that things we have no desire for can benefit us in various ways. For example, you might grudgingly agree to attend the opera with a friend, only as a favor, certain that the experience will be excruciating. Although you have no desire to see the opera *per se* (it is the desire to make your friend happy that motivates you to go), you actually find the opera to be a great experience. We typically form desires with a view to making our lives better in various ways, and to the extent that we dedicate ourselves to the satisfaction of the desires of those we care about, it is usually because we think that this is the best way to improve their lives. However, there is no guarantee that the satisfaction of our desires, or the desires of others, will actually result in a better or happier life. The actual-desire view is unable to account for this very commonplace phenomenon.

A second problem for the actual-desire view is that it makes it too easy for us to improve our lives. If we take seriously the proposition that one's good consists in the satisfaction of one's desires, then we have to conclude that one can improve one's life simply

by altering one's desires so that they conform to those states of affairs which are most easily obtainable. A popular song by Cheryl Crow expresses just such a notion:

My friend the Communist
Holds meetings in his R.V.
I can't afford his gas,
So I'm stuck here watching T.V.

I don't have digital.
I don't have diddly-squat.
It's not getting what you want,
It's wanting what you've got.²¹

No doubt, there is a certain wisdom in such an approach to life. The Epicureans argued that one could improve one's life by ridding oneself of certain desires, particularly those which are neither natural nor necessary – e.g., a desire for political power or social status.²²

Moreover, Sidgwick tells us that a prudent man is accustomed to suppress, with more or less success, desires for what he regards as out of his power to attain by voluntary action – as fine weather, perfect health, great wealth, fame, etc. (110). However, Sidgwick goes on to note that the prudent man who suppresses these desires does not thereby recognize these things to be less good. What is prudent about the suppression of such desires is that their continual frustration will lead to bad states of consciousness. Similarly, the Epicureans claimed that unnatural and unnecessary desires, along with false beliefs, resulted in mental disturbance, and were the main impediment to *ataraxia*. However, the actual-desire view suggests something quite different. It suggests that if I follow Cheryl Crow's advice, and simply train myself to want whatever I already have, or have easy access to, then I will thereby be happier (i.e., my life will go better). This is a deeply counterintuitive result. It suggests that even the

²¹ The title is *Soak up the Sun*.

²² See for example Scholion on Epicurus, in Long and Sedley 21I.

poorest of the poor – e.g., someone living in a garbage dump in a Third-world country – can improve their lives simply by learning not to expect too much out of life.

The inability of the actual-desire view to account for the phenomenon of mistaken desires and its implication that one can improve one's life simply by modifying one's desires are the sorts of problems which have led a number of desire-satisfaction theorists to develop an alternative view. This version of the desire-satisfaction view holds that the good consists in the satisfaction of certain hypothetical desires, usually those desires one would have under certain epistemic conditions. Although there is a long list of philosophers who hold such an informed desire-satisfaction view of the good,²³ I should like to focus on two in particular: Brandt and Peter Railton. These are two of the most influential and sophisticated of the informed desire-satisfaction views, and any critique of subjective theories of the good must address them. Although their views seem to constitute the most plausible and intellectually satisfying versions of subjectivism, nonetheless they are still open to decisive criticisms.

Brandt starts by asking how, and to what extent, desires can be criticized or rationally evaluated (110-112). He claims that a desire is rational or irrational depending on whether it can survive (or be produced by) confrontation with all relevant factual information and logic:

This whole process of confronting desires with relevant information, by repeatedly representing it, in an ideally vivid way, and at an appropriate time, I call *cognitive psychotherapy*. I call it so because the process relies simply upon reflection on available information, without influence by prestige of someone, use of evaluative language, extrinsic reward or punishment, or use of artificially induced feeling-states like relaxation. It is *value-free reflection* (Brandt: 113).

If an agent would still have a particular desire after undergoing cognitive psychotherapy, then

²³ Some examples include Barry, Griffin, Hare, Rawls 1971, and Raz.

the desire is by definition rational; if the desire would extinguish it is irrational. Brandt then identifies an agent's good with the satisfaction of those self-interested desires which would survive cognitive psychotherapy.²⁴ Although it is not completely clear why Brandt insists on a value-free reflection and the exclusion of evaluative language, this seems to be driven by a certain kind of empiricism and a skepticism about normativity. The attempt to deal with moral or ethical problems completely in terms of facts and logic, along with the talk about reforming definitions, gives Brandt's arguments a certain neo-positivistic air.²⁵ However, if Brandt's version of subjectivism is ultimately driven by a general skepticism about normativity, it is hard to see how appeal to facts and logic can be of help, since it would seem that both logic and epistemology are irreducibly normative disciplines.²⁶

Putting aside any meta-ethical concerns about Brandt's methodology, what should we make of his view as a substantive claim about the good?²⁷ I shall argue that Brandt's position

²⁴ See Brandt pp. 113, 127, and 329. Brandt seems surprisingly reticent to explicitly define the good as rational, self-interested desire. Nevertheless, this is how Brandt is usually interpreted—see, for example, Velleman (353) and Sobel (792)—and I will follow this reading for purposes of the present discussion. Part of Brandt's reticence is probably due to his concern about being mistaken as offering a synonymy claim regarding the terms rationally desired and intrinsically good (Brandt: 127). However, I take it that part of the problem is that, although the official view is supposed to be a version of desire satisfaction, Brandt is nevertheless intuitively drawn to something like traditional hedonism, which would identify the good with a type of mental state, such as happiness. See his discussion of the concept of self-interest, with its emphasis on such states as pain, pleasure, enjoyment, and distress (328-331).

²⁵ Note the rather glowing review of Brandt's book by Gilbert Harman, another empiricist who looks at normativity with a skeptical eye (Harman 1982, especially 120-22).

²⁶ This is a controversial claim, and I do not have space here to defend it. I am not relying on this as a central objection against Brandt's position, but it is worth mentioning.

²⁷ Perhaps Brandt would point to the fact that he is offering a reforming definition of the good as evidence that his position does not constitute a substantive theory of the good. However, as Velleman points out, Brandt's claim that we should act in accord with those desires we would have after cognitive psychotherapy, as opposed to the ones we in fact have,

is open to a decisive series of counterexamples; it is simply not plausible to claim that an agent's good consists in the satisfaction of whatever desires the agent would have were she completely informed and logical. Imagine that Suzie is an aspiring physicist. Intuitively, it seems that it would be good for Suzie to learn calculus; however, if Suzie were fully informed, she would not *desire* to learn calculus, since she would already *know* calculus. Thus, it is not the case that for the satisfaction of a desire to be good for an agent, the desire must be one which the agent would have if she were fully informed.

Similarly, it is not the case that an agent will necessarily benefit by the satisfaction of any desire he would have if he were fully informed. Suppose that Johnnie hates the opera. He especially detests Wagner's operas: he doesn't like the music; he doesn't understand German; and they are of a seemingly interminable length. In short, he finds the experience not just unpleasant, but absolutely excruciating. Now, it is plausible to think that if Johnnie were fully informed—where this would include fluency in German, knowledge of the Norse mythology on which the opera is based, and a better general understanding of the art form itself—then he would love the opera, and would desire to go as often as possible. However, the mere fact that a fully-informed Johnnie would desire to attend a Wagnerian opera does not change the fact that our less-informed Johnnie would find the experience excruciating.²⁸

Thus, it is not the case that the satisfaction of any desire that would survive cognitive psychotherapy will benefit us. The case of Suzie shows that we can sometimes benefit from the satisfaction of desires that would extinguish in cognitive psychotherapy, while the case of Johnnie shows that not all the desires that would emerge from cognitive psychotherapy offer a

is a substantive ethical proposal, a proposal about how to live (Velleman: 355).

²⁸ The opera-example is taken from Sher (235n64).

reliable guide to what is good for us as limited, less-than-fully informed agents.²⁹

Allan Gibbard also offers a compelling counterexample against the claim that the good consists in the satisfaction of those desires one would have if fully informed (1990: 20). Imagine that one has a certain neurosis such that if one represents to oneself, fully and vividly, the inner workings of others' digestive systems, then one will never again be able to bear eating among a group of people. Imagine further that the neurosis is such that the *only* way the agent can control it is by avoiding such representations; in other words, the neurosis resists cognitive psychotherapy (the resistance is perhaps deeply embedded in the agent's personality from very early childhood). What would be rational for such an agent? What would constitute the good for such an agent? As Gibbard points out, Brandt's position commits us to the counterintuitive conclusion that this agent ought to avoid eating with others, since a rational desire is *by definition* whatever would survive cognitive psychotherapy. The more intuitively appealing answer is that this agent, given his rather peculiar neurosis, would be better off with less information. In short, Gibbard's counterexample shows that if one combines a certain kind of neurotic psychology with cognitive psychotherapy, then what emerges is something other than the agent's good — i.e., what is intuitively good for the neurotic agent is not necessarily what he would desire after cognitive psychotherapy.

One reason that Gibbard's thought experiment is effective against Brandt's position is that Brandt understands cognitive psychotherapy in purely *causal* terms (Brandt: 113). However things happen to shake out for the individual agent after the process of cognitive

²⁹ This general line of thought can also be found in Sobel (792-93). See also Railton (1986b: 174).

psychotherapy determines what is good or rational for that agent. Thus, an agent with a non-standard psychology may get counterintuitive results. This seems to be a shortcoming in Brandt's account. Moreover, this difficulty is not restricted merely to the neurotic. One can imagine many desires or preferences which are so deeply embedded in our psychology – e.g., the desire to conform to certain social norms – that they would, for many of us, resist cognitive psychotherapy; however, it seems to be an open question whether these desires or preferences are ultimately rational or good for us. Take for example someone who grows up as a slave in a particularly brutal environment. This person may very well go through his entire life with a very strong, deeply-embedded desire or preference to defer to his masters and those like them. Even if the slave gains his freedom, and moves to a place where this sort of deference is neither demanded nor expected, he may nevertheless feel the desire to defer. Moreover, given his experiences in childhood and early adulthood, this desire may not extinguish in cognitive psychotherapy. Brandt's view forces us to call such a desire rational, and to count its satisfaction as contributing to the good of the agent, which again seems seriously counterintuitive.

One might attempt to salvage Brandt's view by insisting on a non-causal understanding of cognitive psychotherapy. Call this version *normative cognitive psychotherapy* (NCP). NCP recognizes that, for some agents, confrontation with logic and all relevant information, even in a full and vivid fashion, is not sufficient to generate intuitively rational desires, or desires the satisfaction of which will necessarily benefit the agent. According to NCP, a desire is rational if and only if it would survive cognitive psychotherapy in a normative sense, that is, in a sense richer than Brandt's causal sense. In other words, NCP would require not only confrontation with all relevant information and logic, but also

any normative or philosophical arguments that might be marshaled.³⁰ Such a view would recognize the neurotic and the slave as engaged in some mistake in reasoning, though not one that is captured in Brandt's conception of logic, or in the mere representation of relevant information.³¹ Thus, NCP does not have to classify the neurotic's desire to eat alone or the slave's desire to defer as rational, or as contributing to the good of the agent. To this extent, NCP is more plausible than Brandt's causal account.

Nevertheless, there are serious problems even with this more normative conception of cognitive psychotherapy. Specifically, there are clear counterexamples to the claim that the satisfaction of any desire that would survive NCP is necessarily good for the agent. As we shall see, these counterexamples trade on the fact that human beings are limited in their power, knowledge, and rationality; moreover, any plausible theory of the human good must take account of these limitations.

To the extent that NCP solves the problem of Gibbard's neurotic or the overly deferential former slave, it is because it recognizes that there are, in principle at least, considerations one could appeal to which could help rid the neurotic of his desire to eat alone or the slave of his deferential habits. In short, the neurosis and the habitual behavior dissolve under NCP. However, given that human agents do not, in fact, always respond to even the best philosophical or normative arguments, the relation between an agent's good and what the agent would desire after NCP will have to be much more complicated than we have

³⁰ Cf. Gibbard's description of reflective equilibrium: To be in reflective equilibrium is roughly to have considered vividly all relevant facts and philosophical arguments, and to have achieved consistent judgments (170).

³¹ The fact that Gibbard refers to this reaction to a vivid representation of digestive systems as a neurosis suggests that there is clearly *something* wrong with the agent's reasoning.

suggested. Gibbard's neurotic, if we send him through NCP, may respond only to enough of the therapy so that he is left with the desire to eat alone, or the neurosis may simply manifest itself in some other undesirable way. The fact that there is, in principle, some philosophical argument such that, if he were to give it fair consideration, it would cure his neurosis altogether, does not change the fact that he does not respond to such arguments, and that he would be better off not representing to himself, fully and vividly, certain facts about the digestive systems of others.

A consideration of weakness-of-will, and the various strategies one might employ in order to overcome it, also shows why our good cannot simply consist in the satisfaction of whatever desires would survive NCP. To say that someone has weakness-of-will is to say that they fail to do that act which they take themselves to have the most reason to do, or they fail to do that which is most in accord with what they take to be their good. One can even say that weakness-of-will occurs when one does something other than what one truly and ultimately desires.³² Let's say that I am a graduate student in the Philosophy Department at UCSD, working on a dissertation having to do with Connectionism, and the eminent connectionist Pat Churchland is my advisor. If I think about it, I realize that part of my reason for wanting to be an academic is that I want to follow in the footsteps of my parents, who themselves were academics. In fact, Professor Churchland reminds me quite a bit of my mother. Moreover, all throughout my life, I have never been able to win my mother's approval. Professor

³² Some would deny this as a conceptual possibility, since they would claim that the strongest or most ultimate desire is the one the agent in fact acts on. However, this seems to be the wrong way to think about strength of desire, at least for our purposes. If we think of our desires as forming a hierarchy, with some playing a more central organizing role in one's life than others, then one can easily imagine being motivated to act by something other than one's strongest and most ultimate desire. I take it that such weakness-of-will is a common occurrence.

Churchland's approval, by contrast, comes fairly easily; moreover, it seems genuine and motivated by an honest appraisal of my work. If I begin NCP, I start to realize that a good deal of my success in my studies rests on a certain cognitive confusion: the main reason I am able to avoid all of the temptations that could distract me from my progress as an aspiring connectionist is that I am under the mistaken impression that through winning the approval of Professor Churchland, I can win the approval of my mother.

Now in the case as I have described it, the desire to please Pat Churchland, to the extent that it depends on a subconscious confusion of her with my mother, will not survive NCP. Many would say that that is for the best; however, I want to argue this is at least not obvious. If my strongest, truest, and most ultimate desire is to be a successful connectionist philosopher, and I also suffer from a fair amount of weakness-of-will, then it may very well be in my best interest to harness certain other (even irrational) desires so that I can achieve what I ultimately want. This is especially the case if there is no reason to think that the irrational desire or belief is harmful in any other way. Let's say that my subconscious confusion of Professor Churchland does not seem to have any adverse effect on my professional relationship with her or other faculty members; also, it doesn't seem to affect my relationship with my mother—that relationship is as good (or as bad) as it would be otherwise. In this case, it seems plausible to conclude that I am better off with the confusion of Professor Churchland with my mother, and the resultant desire to please her, even though these would not survive NCP.

If my ultimate desire is to be a successful connectionist philosopher, one would expect that I would have other reasons—ones which would survive NCP. In principle, I should be able to keep myself going in my studies by rehearsing to myself these normatively

sound reasons. Unfortunately, it is a depressing fact about human nature that normatively sound reasons do not always motivate us to act accordingly. At least some of us, some of the time, may very well be better off with desires or beliefs that would not survive NCP.³³ Thus, an agent's good cannot consist simply in the satisfaction of those desires which would survive cognitive psychotherapy, even if we define cognitive psychotherapy broadly enough so that it includes a consideration of the best philosophical and normative arguments.

This brings us to yet another version of subjectivism—a view associated with Peter Railton (1986a). Railton proposes that an individual's good consists in what he would want himself to want, or to pursue, were he to contemplate his present situation from a standpoint fully and vividly informed about himself and his circumstances, and entirely free of cognitive error or lapses of instrumental rationality (16). It is easy to see the similarities between Railton's view and that of Brandt: they both think of an agent's good in terms of what the agent would think or desire if she were fully informed and free of logical error. The important difference between the two is that, while Brandt identifies the good of an agent with the satisfaction of whatever desires she would have if she were fully informed, Railton defines the good as what a fully-informed version of an agent would want for, or would advise, the agent's less-informed self.

Railton's view seems to avoid all of the problems we have raised for Brandt's version of the desire-satisfaction view. Recall Suzie, our aspiring physicist and Johnnie, the hater of Wagnerian opera. Although the fully informed Suzie has no desire to learn calculus (since

³³ This is especially the case within a subjectivist framework, which I am assuming here for the sake of argument. By contrast, an objectivist about the good may very well deny that irrational desires or false beliefs can ever benefit, since the objectivist usually holds that such things as knowledge and rationality are valuable or good for the agent independently of any desires of the agent.

she already knows it), she may very well desire that her less informed self have a desire to learn it. Similarly, although the fully informed Johnnie has a strong desire to go to the opera, given what he knows about his less informed self, he does not desire that the less informed Johnnie have a desire to go to the opera. The fully informed version of Gibbard's neurotic, assuming that he responds to the best philosophical or normative arguments we might muster against his neurosis and thus is free of the debilitating desire always to eat alone, will tailor his advice to his irrational self according to the real likelihood that his imperfect self will respond to NCP: if the fully informed version of the neurotic has reason to believe that his less informed and less rational self will not ever progress far enough through therapy to free himself of his debilitating desires or beliefs – perhaps because his imperfect self lacks the time or financial resources for the full course of the therapy – then he may simply advise his less informed self not to think about the inner workings of the digestive systems of others. Similarly, the fully informed and perfectly rational connectionist philosopher may very well encourage his less informed and less rational self to continue in the mistaken psychological association of his professional mentor with his mother, especially if he knows that his less informed self suffers from a chronic case of weakness-of-will, and will otherwise, as a matter of fact, never achieve his central goal of becoming a successful connectionist.

Now Railton introduces a potential source of confusion into his account when he speaks of the fully informed self taking the place of the actual self. According to Railton, the wants in question...are wants regarding what [the fully informed version of the agent] would seek were he to assume the place of his actual, incompletely informed and imperfectly rational self, taking into account the changes that self is capable of, the costs of those changes, and so on (1986a: 16). He then says that an individual's *intrinsic* good consists in

attainment of what he would in idealized circumstances want to want for its own sake or, more accurately, to pursue for its own sake (for wanting is only one way of pursuing) were he to assume the place of his actual self (1986a: 17).

There are two ways in which one might interpret this notion of the fully informed agent taking the place of the actual agent. First, one might think that the fully informed agent literally takes the place of the less informed self, retaining all of the characteristics one associates with a fully informed and perfectly rational agent. However, this cannot be the right interpretation, since it would ultimately be indistinguishable from Brandt's view, and all of the problems we have raised for his view would simply reemerge.³⁴ The second and much more plausible reading has the fully informed and perfectly rational agent assuming the place of the actual agent, without retaining his idealizing features. On this interpretation, the fully informed agent asks himself what he would desire for the actual agent, were he to take the place of the actual agent, without retaining full information and perfect rationality.

This notion of the ideal agent taking the place of the actual agent seems to serve two purposes for Railton's view. First, it offers an answer to the question of why we should expect the fully informed version of the agent to be motivated to give good advice to the actual agent. After all, one possible reaction to a full and vivid representation of all relevant information is apathy or indifference toward others.³⁵ Moreover, it is not at all obvious why

³⁴ Note that on this interpretation, I can determine what is rational for me, or conducive to my good, by asking what I would do if I were fully informed and perfectly rational, which is no different from Brandt's position. For example, if the fully informed Johnnie were to replace the actual Johnnie, while maintaining his full information and rationality, then he would want to go to the opera. Thus, on the present interpretation, the actual Johnnie should want to go to the opera, which is counterintuitive given how excruciating actual-Johnnie finds the opera.

³⁵ Consider for example the tendency of some people simply to be desensitized by vivid or explicit representations of violence or the pain and suffering of others.

one should think that benevolence or sympathy is logically or rationally required. However, if the ideal agent must assume the place of his less informed self, then he has at least a self-interested reason to take the best interests of his imperfect self into account. Second, the requirement that the ideal agent take the place of the actual agent renders Railton's account of the good more intuitively appealing for a subjectivist, since it seems to assure that the actual preferences and needs of the agent will be, at least to some extent, taken into account. This is important for Railton, who holds that an important feature of the concept of intrinsic value is that what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive (1986a: 9).

Railton's version of the desire satisfaction view is the most powerful and sophisticated form of subjectivism about the good of which I am aware. It is especially adept at capturing some of the intuitions that lead many to embrace an objective theory of the good. For example, it denies that one's good consists simply in the satisfaction of whatever one happens to desire; it also explains how it is that one can be mistaken about one's good. However, there are serious problems with even Railton's sophisticated version of subjectivism. First of all, in capturing some of our objectivist intuitions about the good, the view lacks many of those characteristics which made more primitive versions of desire satisfaction attractive to many subjectivists. For example, Railton's view, because it rejects the notion that one's good must be transparent to oneself, offers less obvious support for anti-paternalism. Second, the problem of interpersonal and intrapersonal comparisons of utility or well-being, which the desire satisfaction view solved by reducing everything to the common currency of desire satisfaction, now re-emerges, albeit in a different form. In theory, comparisons of utility are not a problem for Railton's view, since we still reduce everything

to the satisfaction of desires or preferences. However, instead of being concerned with the satisfaction of actual desires as revealed by choice behavior, we now have to concern ourselves with hypothetical desires – i.e., the desires one would have under certain counterfactual conditions. It is very difficult to see how the policy maker, who wishes to promote well-being as defined by Railton, or the economist, to the extent that she wishes to predict and explain those forms of behavior which lead to the promotion of well-being, as opposed to mere desire satisfaction, are to do this, given the difficulty in determining what one would desire under these ideal conditions.

By making itself better able to capture some of our objectivist intuitions about the good, Railton's view loses much of its appeal *as a form of subjectivism*. We see this not only in its lessened ability to deal with issues such as paternalism and interpersonal comparison of utility, but also in a certain tension between Railton's account of the good and the internalism which is ultimately supposed to be driving it. Railton has the following to say about internalism:

Is it true that all normative judgments must find an internal resonance in those to whom they are applied? While I do not find this thesis convincing as a claim about all species of normative assessment, it does seem to me to capture an important feature of the concept of intrinsic value to say that what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware. It would be an intolerably alienated conception of someone's good to imagine that it might fail in any such way to engage him (1986a: 9).

It is easy to see how an actual-desire view of the good would satisfy this internalism constraint. What is less clear is how an informed-desire view, such as Brandt's or Railton's, can achieve the same result. After all, according to these views, my good consists not in the

desires that I actually have, but in certain hypothetical desires, desires I would have in certain circumstances. It seems plausible to think that the desires I would have, were I to go through Brandt's cognitive psychotherapy, could be radically different from the desires I now have. Similarly, it would not be at all surprising if my fully informed and perfectly rational self had a set of preferences and dispositions so radically different from my present ones, that I would be very hesitant to view the advice of this 'other self' as a reliable indicator of my good, or even to think of him as a version of myself at all.³⁶ At the very least, it would not be at all surprising if my fully informed and perfectly rational self wanted things for me which failed to engage me in any way, or things which I did not find to any degree compelling or attractive. If parents can want such things for their children, it is not hard to imagine a fully informed and perfectly rational self wanting the same for a less informed and imperfectly rational self.

The analogy of parents making decisions for their children is instructive. Typically, we think that parents are in a position to make authoritative judgments about what is good for their children because the children themselves lack the knowledge and practical reasoning skills they would need in order competently to make those decisions for themselves. Might we say the same about the fully informed advisor and her actual self? I think we would only if we thought that the actual self had some reason to think that she was growing or developing into something like the fully informed and perfectly rational self, much in the same way that most children are on the way to becoming normal, better informed and more rational adults. Unfortunately, our nature is such that we will always be less than fully informed and less than

³⁶ Connie Rosati argues, persuasively I think, that we have reason to doubt that the judgment of a fully informed agent is something that we should find authoritative. Thus, she claims that Railton's view lacks normative force (Rosati: 299, 307-314).

perfectly rational. This is part of who we are, and any plausible theory of the good must take this into consideration. For my part, I find it difficult if not impossible to conceive of a fully informed and perfectly rational version of myself. Any such being would be so different from me as to seem utterly alien. It would not be a version of me, and I see no reason to think that its judgments about my good would be reliable or authoritative.

Conclusion

In this chapter I have considered so-called subjective conceptions of the good. As we have seen, with regard to certain versions of hedonism this label can be a misnomer since there are versions of hedonism which are most sensibly classified as objective accounts of the good. Henry Sidgwick's account of the good as desirable consciousness and the Epicureans' identification of the good with *ataraxia* or tranquility are examples of such objective versions of hedonism. But even though these views are ultimately objective since they do not make the good contingent on agents' desires, choices, preferences, or beliefs, they are at least subjective in the weaker sense of taking the good to consist in certain kinds of subjective experiential states. Nozick's experience-machine thought experiment shows all such accounts of the good to be indefensible. Other versions of hedonism are subjectivist in the stronger sense of taking the good to consist in certain types of mental states *because of or in virtue of the fact* that such states are ultimately what we desire. Such views are ultimately indistinguishable from hedonism's main subjectivist rival, the desire-satisfaction account of the good.

I considered two versions of the desire-satisfaction view—the actual-desire account and the informed-desire account. The actual-desire account is clearly the less plausible of the

two, due to the rather commonplace phenomenon of harmful desire satisfaction and beneficial desire frustration. The recognition of this widespread phenomenon has led some desire-satisfaction theorists to propose more sophisticated versions of the view. Richard Brandt identifies the good with the satisfaction of those desires one would have under conditions of full information and complete rationality. Peter Railton, by contrast, identifies the good of an agent with what a fully informed and fully rational version of the agent would want for her less knowledgeable and imperfectly rational self. Although these informed-desire versions of the account are more plausible than the actual-desire view, they are still ultimately indefensible. Brandt's view still has some extremely implausible implications about the good. For example, it suggests that, since a fully informed version of an engineering student has no desire to learn calculus since she already knows calculus, the student herself has no reason to learn calculus. Railton's account avoids some of these problems by stipulating that the fully informed version of the agent is to advise or even take the place of her less informed self; however, this has the effect of putting a significant amount of psychological distance between the two selves, and it is not at all clear that the preferences of the fully informed self are a normatively sound guide to the good of an imperfectly rational agent with limited knowledge.

It is also important to see that the sorts of views held by Brandt and Railton offer little in the way of practical guidance about the good. They both recognize that one's good does not consist merely in the satisfaction of one's present, actual desires. In that way, their views are similar to objective account of the good. However, it is not at all clear that either Brandt or Railton has much to say about what particular judgments about the good would survive cognitive psychotherapy or be suggested by our fully informed and perfectly rational selves. In order to give a plausible answer to such questions, we cannot avoid asking difficult

questions about the human good, and it does not seem that the informed desire-satisfaction view can be of much help in this quest. The question of what we ought to desire, or what we would desire if we were fully rational and fully informed, seems to be at heart a normative question; specifically, it is a question about the human good. In the next chapter, I will turn to an attempt to answer this question in terms of the perfection of human nature.

Chapter 3 Hurka's Perfectionism

I. Introduction

In the previous chapter, I argued that subjectivist accounts of the good, both in their hedonistic and desire-based forms, are untenable. But if we are not to be subjectivists about the good, then what sort of account should we accept? Clearly, any adequate account must avoid the sort of objections which undermine subjectivist views. For instance, our account must not be vulnerable to Nozick's experience-machine thought experiment. Also, it must not have the sort of counterintuitive and ultimately implausible implications which plague various versions of the desire-satisfaction view. Perfectionism offers an interesting and important alternative to these subjectivist accounts. It is objectivist because it denies that the good depends on the desires, beliefs, preferences, or commitments of agents. Since perfectionism conceives of the good in terms of the development and exercise of certain kinds of capacities, as opposed to the possession of certain kinds of experiential states, the view is immune to concerns raised by Nozick's experience machine. Moreover, since perfectionism sees the good in objective terms, and does not take the good of an individual to depend on her actual or hypothetical desires, the view promises to offer more in the way of *guidance* about the good than do desire-based accounts. For instance, compared to Peter Railton's informed desire-satisfaction account of the good, the objectivism of perfectionism allows us to say something more concrete and specific about what a fully rational and fully informed agent would desire.

In this chapter, I consider a contemporary version of perfectionism recently offered by Thomas Hurka. Hurka describes his position as Aristotelian, and contends that the human

good consists in the development of those properties which are essential to human nature. Hurka's view is interesting and important for several reasons. First, it is a decidedly objectivist account of the human good, insisting that certain things are good for us independently of our beliefs, attitudes, desires, or responses, and thus represents a clear alternative to subjectivist accounts. Second, although Hurka's view has some things in common with Aristotle's position, he makes every effort to avoid the sort of controversial metaphysical commitments one usually associates with Aristotelianism. For example, Hurka makes no appeal to a teleological worldview, or to Aristotle's related doctrine of a proper human function. Instead, in establishing his claims about the human essence, Hurka restricts himself to a purely descriptive concept of human nature, along with the type of essentialism associated with Hilary Putnam and Saul Kripke.

Given the widespread acceptance of Putnam/Kripke-style essentialism—at least in the contemporary Anglo-American tradition—Hurka's project, if successful, would represent a significant philosophical achievement. Indeed, Hurka has been praised for making perfectionism a live option in contemporary ethical theory, a serious alternative to more familiar consequentialist, deontological, and virtue-based accounts of ethics. According to Tim Mulgan, Hurka ably demonstrates that the application of modern analytical techniques can transform perfectionism into a serious rival to contemporary Kantian, Utilitarian, and Virtue Theories (550). L.W. Sumner, meanwhile, has the following to say about Hurka's treatment of the view in *Perfectionism*: Long neglected by analytic philosophers, perhaps because of its suspect links with idealists (Bradley, Green, Bosanquet) and crazies (Nietzsche), its reputation should undergo a thorough rehabilitation as a result of Tom Hurka's delightful and provocative book (1995: 151). Similarly, although he is critical of

certain aspects of Hurka's argument, Thomas Carson says that Hurka has considerably expanded the range of standard views that philosophers need to consider (723).

Although I agree with these reviewers that Hurka's book is clever and resourceful, and that he probably does succeed in his central aim of elevating perfectionism to the status of a serious player in contemporary ethical theory, I also think that the view, as Hurka presents it, is ultimately untenable. Hurka's version of perfectionism suffers from two major defects. First, it is committed to implausible claims about the human essence. Hurka claims that the human essence consists in physical perfection and rationality; however, we have good reason to think that such characteristics are not essential to human beings. Second, his view offers an impoverished and unattractive account of the human good; even if Hurka were right about the human essence, the claim that the human good consists in the development of that essence (as Hurka understands it) would still be implausible and deeply unattractive since it leaves out other important goods such as pleasure, happiness, contentment, and evils such as pain, depression, and anxiety. In short, while I sympathize with Hurka's objectivist inclinations and share his suspicion of and dissatisfaction with subjectivist accounts of the good, his perfectionist account is one which we ultimately cannot accept.

II. Hurka's Perfectionist Account

I shall begin by giving a brief sketch of Hurka's account of the good. On Hurka's view, the human good is grounded in human nature; that is, the human good consists in the development of certain characteristics which are essential to human beings and conditioned on their being living things (1993: 17). According to Hurka, there are three aspects of the human good understood in perfectionist terms: physical perfection (or the development of our

physical nature), theoretical rationality, and practical rationality (37). For Hurka, what is good is the development of these three characteristics to a high degree. Physical perfection, for Hurka, ends up being a matter of health and physical fitness, so that the Olympic athlete best exemplifies this sort of perfection (39). Theoretical rationality involves the exercise of those capacities one typically associates with the great scientist, whereas practical rationality is a matter of the successful pursuit of a complex and suitably extended network of goals, and is best exemplified by the successful political leader. As a way of illustrating these concepts, one might say that the ideal life³⁷ would be one which combined the body and athletic achievements of Charles Atlas,³⁸ the theoretical understanding of Albert Einstein, and the

³⁷ Even though Hurka argues for an agent-neutral, consequentialist form of perfectionism, he nevertheless thinks that, in aggregating units of perfection, we should start with the value of particular lives. More precisely, he argues that a consequentialist perfectionism should aggregate first across times and only then across persons. It first calculates the aggregate good in each life considered as a self-contained unit and then combines these measures across lives. The alternatives aggregating first across persons at a time or simultaneously across persons and times would abandon perfectionism's traditional focus on the life as a morally significant unit (1993: 69). Thus, although Hurka seems to echo G.E. Moore's rejection of talk of the personal good as nonsensical (see especially pp. 17-18, where Hurka says that his perfectionism gives an account of what is good *in* a human life and makes no claims about what is good *for* a human, and denies that there is conceptual room for the notion of perfectionist goods being good *for* the agent who exemplifies them), he nevertheless thinks that it is crucial that we begin the process of aggregation by first adding up the units of perfection within lives. This is partly because we cannot know the value of a given perfectionist state or accomplishment without first knowing facts about the life in which it occurs. For example, four moderately successful years as the mayor of a small town in Iowa would offer more units of perfection to the body builder Atlas or the physicist Einstein than it would the towering political figure of FDR. By contrast, a moderate physical accomplishment, such as becoming an above-average golfer, would give more perfectionist units to FDR or Einstein than to Atlas. On Hurka's view, while we can make sense of the claim that Atlas's life is good, we cannot make sense of the claim that Atlas's life is good for Atlas; he simply exemplifies or realizes a high degree of physical perfection. For a discussion of what counts as a well rounded life for Hurka, see chapter 7 of *Perfectionism*, along with Hurka 1987. For Moore's claims about the personal good, see pp. 96-102 of *Principia Ethica*.

³⁸ It is worth noting that, with his emphasis on the importance of running and jumping (1993: 39), Hurka might disapprove of Atlas as the paradigm of physical perfection, instead

practical accomplishments of a political leader such as Franklin D. Roosevelt.

As Hurka points out, perfectionism, understood as the view that the good consists in a certain kind of excellence defined in terms of human nature, has a long and impressive history. Its adherents have included Plato, Aristotle, Aquinas, Spinoza, Leibniz, Kant, Hegel, Marx, Nietzsche, Green, Bradley, and Bosanquet (1993: 3-4). However, perfectionism fell out of favor some time in the early part of the twentieth century, and has been all but ignored by philosophers since about 1920 (Hurka 1987: 727). According to Hurka, part of the problem with perfectionism is that it has come to be associated with a number of discredited, or at least highly contentious, philosophical doctrines. Examples of such doctrines include the following: (1) the notion, associated with philosophers such as Aquinas and Spinoza, that goodness and being (or alternatively, perfection and reality) are the same, and that an increase in goodness (or perfection) constitutes an increase in being (or reality); (2) notions of positive freedom held by philosophers such as Green and Bradley, who claim that true freedom consists in the choice of those options that most develop our nature or realize our true selves; (3) a teleological worldview, along with the related notion that human beings have a proper function; (4) the idea that the good for humans consists in the development of those characteristics that are *distinctive* of them or *unique* to them as a species or kind; (5) various sorts of natural tendency doctrines, according to which human beings have a natural tendency or desire to develop their essential natures to a high degree (1993: 23-26). Hurka

preferring someone like Jesse Owens or Johnny Weissmuller. However, given that Atlas was named the world's most perfectly developed man in a physical fitness exhibition at Madison Square Garden in 1922, a title which he used as a sort of personal and professional slogan for the rest of his life, the temptation to use Atlas is too strong to resist. According to an Atlas-related web site, his measurements are stored in a time capsule at Oglethorpe University.

thinks that all of these doctrines are controversial or problematic enough to act as bars to the acceptance of a perfectionist morality. However, he also thinks that none of them is essential to perfectionism; he calls them accretions and claims that perfectionism is better off without them. In fact, Hurka offers a stripped down version of perfectionism, which avoids the sorts of metaphysical and psychological doctrines many have found so problematic, as the best and most plausible version of the view.

Hurka's identification of the good with physical perfection and rationality is certainly an interesting and controversial position. But how is he to support such claims if he rejects the sorts of metaphysical doctrines which have traditionally undergirded perfectionist views? Hurka makes use of what he takes to be purely descriptive, non-teleological concepts of human nature and human being, along with the sort of essentialism made popular in the last thirty years or so by Putnam and Kripke. Hurka thinks that these tools, along with a Rawlsian-inspired approach to the justification of moral claims,³⁹ are sufficient to establish his identification of the human good with states such as theoretical rationality, practical rationality, and physical and athletic excellence; he argues that no other metaphysically suspect doctrines are necessary for his perfectionism.

Hurka has two stories to tell if he is to adequately support his perfectionist claims about the good. First, he must convince us that physical perfection, theoretical rationality, and practical rationality, as he understands these concepts, are part of the human essence, and that their development and exercise constitutes a development of our nature or our humanity. Second, he must convince us that the human *good* consists in the development and exercise of

³⁹ For the Rawlsian account of the justification of ethical claims see Rawls 1951 and 1971 (especially 46-53). See also Daniels 1979 and 1980.

these purported essential traits. It is the first story that involves an appeal to Putnam/Kripke-style essentialism, and although this view is complex, for our purposes here a truncated description of it will suffice.⁴⁰ Both Putnam and Kripke think that natural kinds, such as water and gold, have an essence. The essence of water is H₂O, while the essence of gold is atomic number 79; anything that lacks atomic number 79 is not gold, no matter how much it resembles gold, and anything that is not H₂O is not really water, no matter how much it shares the phenomenal or surface characteristics of water. It is through a combination of intuitive judgments and empirical research that we discover such essences, so that although the essence of water has always been H₂O, human beings did not become aware of this fact until sometime within the last three hundred years or so. In his explanation of how the human essence is known, Hurka argues that we can apply such an approach to human nature, and when we do, we find that our intuitive judgments, along with empirical sciences such as biology and cognitive psychology, reveal that physical perfection and rationality are indeed essential to human beings (1993: 33-44).

What about the second story? That is, how does Hurka connect the notion of a human essence to the human good? Hurka thinks there are various considerations supporting the claim that the good consists in the development of the human essence. First, he thinks that the notion that the human good consists in developing human nature — i.e., that our good consists in becoming more fully human — is a deeply attractive and intuitively plausible idea. Second, the number and diversity of famous philosophers who have been attracted to such an ideal is, Hurka thinks, a mark in favor of perfectionism. Third, and most importantly, Hurka

⁴⁰ Kripke's arguments regarding essentialism appear in *Naming and Necessity*, while Putnam's views are to be found mainly in "The Meaning of Meaning" and "Is Semantics Possible?" I discuss their views in a little more detail in section IV.

argues that his perfectionist account of the good can *explain why* we think that certain things such as rationality and athletic excellence are valuable even if certain people do not value them. Moreover, he thinks that his view can offer a unified account of a wide range of judgments we make about the good, the right, justice, and so forth. Perfectionism's promise of explaining and unifying our various considered judgments about such questions is perhaps the most attractive feature of Hurka's view, and if he is successful, then he will have gone a long way toward establishing the truth of perfectionism.

However, I shall argue that neither of Hurka's two stories is very believable. We have good reason to think that physical perfection and rationality are *not* essential to human beings. Indeed, it is extremely doubtful that human beings have any essence at all. Second, I shall argue that Hurka fails in his attempt to give a unified account of the good; important goods are left out, and at the very least his account of the good is strikingly incomplete. Moreover, if Hurka attempts to remedy this incompleteness problem by adding subjective goods such as pleasure and enjoyment to the list of goods, without giving some perfectionist account of the value of such states, then his claim that perfectionism unifies and explains our particular ethical judgments is seriously undermined. But before I turn to these criticisms of Hurka's position, I want to discuss some criticisms of Hurka's view recently made by Philip Kitcher. Although I am sympathetic with much of what Kitcher says about Hurka's view, ultimately Kitcher leans too heavily on certain methodological assumptions which, due to their own defects, undermine his critique.

III. Kitcher's Critique of Hurka

In *Essence and Perfection*, Kitcher notwithstanding Hurka's explicit rejection of

much of the traditional metaphysical baggage of perfectionism attacks Hurka's view at the foundational level. Specifically, Kitcher thinks that Hurka fails to give an adequate explanation for why it is that certain states are valuable; in other words, Hurka fails to meet what Kitcher calls the Reductivist Challenge. I want to claim that Kitcher is unfair to Hurka on this score, but to see why we need to look more closely at Kitcher's central complaint. Kitcher starts by distinguishing two types of objectivism about the good (59). The first type, *bare objectivism*, simply offers a list of the things that make a life go well, with no explanation of how these things get on the list. *Explanatory objectivism*, by contrast, attempts to explain and justify the content of the list. As Kitcher rightly points out, the long history of widespread disagreement about the human good makes bare objectivism seem unattractive from both practical and theoretical points of view. Surely, an adequate account of the good should have *something* to say about why certain states or activities are intrinsically good while others are not. Hurka's perfectionism, with its appeal to the human essence and its claim that the good consists in the development of that essence to a high degree, is a version of explanatory objectivism.

So far, there is nothing especially controversial or objectionable in Kitcher's analysis of Hurka's view. However, problems arise when Kitcher specifies what he takes to constitute an adequate justification or explanation of an objectivist claim. According to Kitcher, if explanatory objectivists are to avoid the sorts of intractable disagreements and immediate stalemates one would expect to arise among bare objectivists,

then it appears that their explanation and justification must take a very particular form. For suppose that the contents of the list are to be grounded in two principles, one that identifies human lives as going well insofar as they exhibit a certain generic property and one that connects the items alleged to be objectively valuable with the favored property.

If the property in question can only be attributed by already making a judgment about what is valuable, then just the controversies that were supposed to be avoided will recur at the higher level. Thus, to improve on bare objectivism, explanatory objectivism apparently must pick out some property whose ascription *can be made in a value-free fashion*, seeing this as the criterion of human well-being (60; emphasis added).

This is Kitcher's statement of the Reductivist Challenge, and it would be helpful to spell out exactly how Hurka's view might be thought to violate it. Let's say that the items on the list of goods are knowledge, friendship, and physical fitness. Presumably, the principle that identifies human lives as going well insofar as they exhibit a certain generic property is Hurka's principle that says that human lives go well insofar as they exhibit or exemplify the development of human nature or the human essence to a high degree (human nature or the human essence being the certain generic property). The second principle, the one that connects the items alleged to be objectively valuable with the favored property is the principle that says that knowledge, friendship, and fitness are part of the development and exercise of human nature or the human essence. Most of Kitcher's argument consists in the denial that the notions of human nature or the human essence—as Hurka understands them—can be given a value-free specification.⁴¹ States such as knowledge, friendship, and fitness will be thought to be tied to human nature or essence only because one already judges them to be valuable. Since the notion of human nature or essence is itself value-laden, the view that identifies the good with the development of human nature or essence fails to satisfy the demands of the Reductivist Challenge.

⁴¹ Kitcher is especially critical of Hurka's claim that rationality is part of the human essence (see especially 71-76). See 68-71 for Kitcher's criticisms of Hurka's claims about physical perfection and the human essence.

What are we to make of Kitcher's Reductivist Challenge and his claim that Hurka fails to meet it? In one sense, this seems undeniably true. For Hurka is clear in his insistence that the concept of human nature be understood as a purely descriptive category (1993: 18-21). However, we should note that this is not sufficient to make Hurka a *reductivist*. Indeed, Hurka makes clear that his main reason for insisting on a descriptive concept of human nature is his desire to avoid moralistic forms of perfectionism. According to such views, human nature is at least partly an evaluative concept, and realizing one's humanity involves the exercise of moral virtue or acting in accord with the demands of morality, where this is often taken to involve following the rules of conventional morality. Hurka associates such moralistic forms of perfectionism with Aristotle, Aquinas, T.H. Green and, in particular, Kant, and thinks that they are misguided (19-20). Hurka argues not only that such views are objectionably circular accounts of morality, but also that there is no reason to think that virtue or the disposition to act in accord with the demands of morality is essential to human nature. Hurka tells us that the kind of rationality that can plausibly be thought of as essential to human nature i.e., means-end reasoning is as easily exemplified by a successful burglar as by a philanthropist.

Regardless of his skepticism about moralistic versions of perfectionism, it's clear that Hurka is not opposed to the appeal to values *per se*: A principle that uses values to identify the morally important properties of humans will be undermined only if the values it uses are moral (18-19). Hurka also says that not just any evaluative content would undermine perfectionism. Only *morally* evaluative content would have this effect (18). One might object that Hurka does indeed appeal to moral judgments in picking out physical perfection and theoretical and practical rationality as essential to human nature, with the idea being that

he favors these properties for moral reasons. However, we should notice that Hurka is using the term *moral* in a familiar but restricted sense. In this context, morality is to be distinguished from ethics, where the former refers to familiar other-regarding duties and virtues, while the latter is a broader term encompassing both our ideas of other-regarding moral duty and our judgments about value and what constitutes a good life.⁴² Traditional eudaimonistic ethical theorists such as Aristotle have held that familiar other-regarding virtues are essential to human nature, while a moralistic perfectionist such as Kant insists that the development and exercise of our rational agency capacity requires the fulfillment of other-regarding duty since this is a rational requirement of the moral law. Thus, moralistic versions of perfectionism make morality essential to human nature, and Hurka thinks that such claims about human nature are simply implausible.

Hurka, since he is a consequentialist, appeals to agent-neutrality to generate other-regarding moral reasons. However, this does not mean that Hurka thinks it illegitimate to appeal to moral judgments in order to establish what in fact constitute our moral obligations.⁴³ Indeed, the strategy of Rawlsian reflective equilibrium *requires* that one make at least some appeal to such judgments; without such an appeal there will be no moral intuitions to balance, modify, and systematize. Similarly, it is not at all clear why it would be absolutely illegitimate to appeal to our pre-theoretical judgments about the human good in determining which properties or characteristics are essential to human nature. Admittedly, such an

⁴² See pp. 30 and 41 of *Perfectionism*. For an influential statement of the relation between ethics and morality see Williams, *Ethics and the Limits of Philosophy* and *Morality: An Introduction to Ethics*.

⁴³ See for example Hurka's arguments for agent neutrality (62-63) where he argues for agent-neutrality as the best account of our other-regarding duties due to its ability to capture our particular judgments about moral obligation.

approach would yield a normative, value-laden conception of human nature, and this is something that Hurka is *in fact* committed to avoiding. However, given his stated reasons for preferring a purely descriptive conception of human nature, and his tendency to recognize moralistic conceptions of human nature as the only real alternative to descriptive conceptions, it is not at all clear that Hurka must embrace that part of Kitcher's Reductivist Challenge that insists that the properties identified as essential to human nature must be picked out without any appeal to judgments about the human good. Such an appeal to our pre-theoretical judgments about the good is also rendered less problematic by the role in Hurka's account of descriptive sciences such as biology and cognitive psychology in identifying those properties which are essential to human nature, since this makes it clear that, in identifying the human essence, we cannot merely pick out whichever properties we are inclined to approve of; any candidate property must also plausibly be attributable to all members of the class *human being*.

Kitcher, by contrast, seems to be motivated by more general concerns about the metaphysical and epistemological status of value claims. This comes out most clearly in the final section of Kitcher's article (78-83). Near the beginning of that last section, Kitcher repeats his claim that we need to move beyond bare objectivism if we are to avoid dogmatic foot-stamping in our attempts to resolve moral disagreements. But then he goes on to say the following:

Moreover, absent some insight into the ways in which the justification [of explanatory objectivism] works, the objectivist metaphysics will seem suspect in just the way that Platonic metaphysics of mathematics seems suspect; it is not that value of the items on the list is questioned, any more than the truth of standard arithmetic is doubted, but that the assertion of objectivity raises the question of how we could ever be justified in making claims about what is valuable (as

the thesis that arithmetic is about abstract objects raises concerns about how we could ever fathom the properties of those objects) (79).

Here, when Kitcher wonders how we could ever be justified in making claims about what is valuable, we seem to have a version of the familiar metaphysical question of how it is that there can be value or normativity in the natural world.

At one point, Kitcher considers the possibility that the objectivist about value could claim that what is needed is a *justification* and not an *explanation* of value judgments. The idea here is that justification and explanation are distinct, and that the practices of pre-Newtonian astronomers such as Kepler, who was very successful in predicting, tracking, and computing the orbits and trajectories of the planets without being able to explain why they moved in the way they did, offer real examples from the history of science of justification without explanation. What is most interesting about this proposed analogy for our purposes is Kitcher's response to it:

Could objectivists make use of this analogy to forswear the need for an explanatory account of the contents of their list? Only if they could point to some way of making the particular judgments about what is objectively valuable. So we might suppose that, as we consider the lives of our friends, of historical figures, of characters in fiction, we just see (or intuit) that some things contribute to their flourishing. Now even if we were to concede that this is the sort of activity in which people standardly engage, serious questions about the status of what is going on would remain. Why not suppose that we are detecting the factors that contribute to human pleasure, or that we are expressing our own subjective desires, or that we are merely recapitulating the styles of judgment instilled in social conditioning? Why suppose, in particular, that we are identifying a kind of good that is independent of our desires? (80)

Here, Kitcher seems to be echoing familiar concerns about the epistemic status of our moral

judgments.⁴⁴ He virtually ignores the possibility that the justification-without-explanation approach to ethics could take place along the lines of Rawlsian-style reflective equilibrium.⁴⁵ On such an approach, we could appeal to our intuitions about what is valuable in various contexts, without making any assumptions about these intuitions offering any mysterious access to occult-like moral properties. On the Rawlsian approach, these intuitions need be thought of as nothing more than our pre-theoretical judgments about what is valuable. The justification of particular judgments about, say, the good, merely consists in working out the most plausible and coherent system of the relevant intuitions. According to Kitcher, however, if we are going to appeal to our intuitions about what is valuable, we need a theory about what the grounds of value are so that when we trade intuitions we can have some confidence that we're expressing the sort of knowledge we're supposed to have (80).

When discussing the Reductivist Challenge, Kitcher makes it clear that it is value judgments in general, and not just judgments about the human good in particular, that he thinks of as metaphysically and epistemically problematic. Going back to the first statement of the challenge, note that Kitcher says that an acceptable explanatory objectivism must pick out some property whose ascription can be made in a value-free fashion (60). In the final section of the article, he says that to meet the Reductivist Challenge one would have to

⁴⁴ See Mackie for a clear and influential statement of the sorts of metaphysical and epistemological concerns which seem to animate Kitcher.

⁴⁵ Kitcher does briefly consider a Rawlsian-type coherentist view about the justification of value judgments as an alternative to his reductivist, foundationalist view (82). While he does seem to back off a bit from his reductivism, saying that we cannot know in advance if the Reductivist Challenge (or something tantamount to it) will prove sound, in the same breath Kitcher says that the prospects for objectivism ought to be assessed by taking up the hard questions of the epistemology of value (82). I take this as simply a restatement of Kitcher's general skepticism regarding the metaphysical and epistemic status of value judgments, and his conviction that reductivism can be avoided only if we can first develop a more attractive or plausible or compelling account of the place of value in the natural world.

specify a criterion of centrality in a way that was uncontaminated by judgments of value (80). Finally, in the last paragraph of the essay, Kitcher claims that what is significant about Hurka's appeal to the notion of a human essence is that, if it would have been successful, then he would have been able to sidestep hard questions in moral epistemology (82). Kitcher also says that the cogency of the Reductivist Challenge involves epistemological questions that are as daunting in ethics as they are in the philosophy of mathematics (83). In short, whereas Hurka is drawn to a descriptive concept of human nature and the human essence because he wants to avoid offering an account of the good and morality that is objectionably circular or simply implausible, Kitcher seems to be driven by a general skepticism or distrust of value judgments.⁴⁶

Although it is clear that Hurka is not interested in meeting the Reductivist Challenge, his discussion of this issue is potentially misleading since, as far as I can tell, he never actually uses the term reductivism. Instead, Hurka talks about naturalism, which he understands fairly narrowly:

Many philosophers assume that a successful defense of perfectionism will involve *naturalism*, in the meta-ethical sense of naturalism. It will not justify moral claims by connecting them to other moral claims, either more particular or more general. Instead, it will derive them from non-moral facts. It will argue that facts about human nature directly entail conclusions about the human good. Thus, if it is essential for humans to be rational, it follows logically that humans ought to develop rationality (1993: 28).

What's potentially misleading about this discussion is that what Hurka is calling naturalism is probably more accurately referred to as analytical reductivism or analytical reductive

⁴⁶ See Brandt (16-23) for an instance of another reductivist who is driven by general metaphysical and epistemological qualms about value judgments and the appeal to ethical intuitions.

naturalism. Broadly speaking, naturalism is the view that all that exists are those entities or states countenanced by the natural and social sciences, or that the proper method of inquiry is that of such sciences.⁴⁷ By naturalism Hurka seems to have in mind the sort of reductive ethical naturalism that G.E. Moore attacked with his open question argument i.e., the sort of ethical naturalism associated with people such as Herbert Spencer. Hurka rejects this version of ethical naturalism, and calls himself an anti-naturalist (1993: 28). However, he could just as easily have referred to himself as a non-reductive naturalist and perhaps *should* have in order to avoid confusion.

Kitcher, with his Reductivist Challenge, insists that the perfectionist who understands the human good in terms of human nature and its essence explain how we can go from facts about human nature to conclusions about the human good without appeal to value judgments, and he accuses Hurka of trying to sidestep difficult questions of moral epistemology. However, it is clear that this is not what Hurka is trying to do. He rejects reductive ethical naturalism, and explicitly denies that moral claims can be understood in purely descriptive, non-moral terms (1993: 28, 30). According to Hurka, the perfectionist ideal must be defended not by logic, but by substantive moral argument (30). Moreover, he says that, on his view, perfectionism is not a magical entrée into morality, but a substantive position within it (30). While Hurka does think that there is a fact of the matter regarding what

⁴⁷ The latter claim is often called *methodological naturalism* while the former is referred to as *ontological naturalism*. Most naturalists seem to hold both of these views. At any rate, both are distinct from any claims about moral judgments or propositions *logically following from* or *meaning the same as* certain non-moral judgments or propositions. Moreover, the same goes for other properties, entities, and terms that seem to resist reduction to a basic science such as physics, including color, meaning, and intentionality. See Schmitt, Lacey, and Post for brief but helpful discussions of naturalism. See also Brink 1989 (9) and Smith (17).

constitutes human nature, conclusions about the good follow from premises about human nature only given the further evaluative premise that developing whatever constitutes human nature is good (1999: 47).

This further evaluative premise is supported by various considerations, some of which I have already touched on. First, the notion that the human good consists in the development of what is essential to human nature is, Hurka thinks, intuitively appealing and plausible. Second, there is the long tradition of perfectionist views in the history of philosophy. Third, Hurka thinks that the human-nature view allows us to explain, justify, and unify a wide range of our intuitions regarding the good, the right, and justice (1993: 31-33). Also, there is the middle section of Hurka's book (55-143), where he offers what Kitcher himself describes as an interesting, subtle, and valuable discussion of questions regarding the measurement, aggregation, and balancing of perfectionist goods (Kitcher: 61). These are all components of Hurka's argument for the substantive moral claim that the good consists in the development of human nature. While it remains a legitimate question at this point how compelling or persuasive these arguments are, Hurka has at least offered an impressive example of how moral justification and explanation can proceed without reductivism, and without falling back on bare objectivism.

Hurka's metaethical assumptions seem very much within the mainstream of contemporary ethical theory. By contrast, Kitcher's reliance on the Reductivist Challenge, with its attendant obsession with an alleged fact/value gap, seems to betray a striking neo-positivist orientation. Although such a conception of philosophy in general, and ethics in particular, could be the right view, it is nevertheless a fairly dated view, and Hurka seems to be well within his rights to reject such an approach. For Hurka, the claim that the good

consists in the development of the human essence is a substantive claim about morality, and must be evaluated on its merits. We have to ask, first, whether the properties favored by Hurka are indeed essential to human nature, and second, how plausible it is to think that the good consists in the development and exercise of those properties. These are the subjects of the next two sections.

IV. The Human Essence

As we saw at the end of section II, Hurka appeals to the Putnam/Kripke story about essentialism in support of his claims about the human essence. Hurka identifies two distinct methods for discovering essential properties; one method involves the appeal to our intuitions about essences, while the other appeals to scientific explanations. Here is what Hurka says about the two methods:

The first method, associated with Kripke, is *intuitive*. It says that we discover essential properties by making intuitive judgments in thought experiments involving candidate members of a kind. To learn whether its atomic structure is essential to gold, for example, we imagine a series of possible substances with gold's atomic structure but a different outward appearance. If we judge all these substances to be gold, our judgments show that its inner constitution is essential to gold and its phenomenal properties contingent. We learn what could and could not be gold by asking how we could and could not imagine gold's existing.

The second method examines *scientific explanations*. Associated with Hilary Putnam, it says that we identify essential properties by their essential role in the explanations given by good scientific theories. That gold has a certain atomic structure explains its colour, weight, and other phenomenal properties, but is not in turn explained by them. Gold's atomic structure is thus explanatorily prior to these properties, and this shows, on the second view, that it is essential. For the explanatory view, properties essential to scientific explanations are essential, and in gold these cluster

around its inner constitution (1993: 34).

Now the first thing to note is that, as a statement of the methodologies of Putnam and Kripke, this does not seem entirely accurate. Although in arguing for his essentialist claims, Kripke does often appeal to his reader's intuitions about hypothetical cases, while Putnam, by contrast, emphasizes reliance on empirical investigation, it seems wrong to suggest that the two endorse differing methods for coming to know essences. Indeed, both Putnam and Kripke seem to assume that, in order to know the essential properties of an object, we must rely on a combination of empirical investigation and appeal to intuition. For example, Kripke could not have thought that we can know the essence of gold *merely* through the appeal to our intuitions about hypothetical cases. At best, such a method would reveal, not the *real essence* of an object, but instead what Locke would call a *nominal essence*, or what contemporary philosophers would call a Fregean *sense*; that is, such an appeal to intuition would merely reveal those properties we conventionally associate with the object in question, without necessarily telling us anything about the true nature of the object. At any rate, Hurka himself suggests that we should use both methods simultaneously, testing our claims about the human essence both intuitively and against explanatory theories (1993: 35). Thus, in order to establish the conclusion that human beings are, say, essentially rational, Hurka will have to appeal to the relevant empirical sciences.

It is clear that Hurka understands *human being* as a biological concept. Indeed, on page 45 of *Perfectionism* Hurka insists that *human* is a biological concept, and on page 46 he appeals explicitly to the category of *homo sapiens*. Also, in discussing both physical perfection and rationality, Hurka says that beings which did not share certain characteristics with us (a certain sort of physical form in the case of physical perfection and a certain kind of

mental life in the case of rationality) would not be part of our species (1993: 37, 39). Since Hurka wants to understand human being as a biological category, then it seems that the most relevant empirical science to appeal to is biology. Thus, in evaluating Hurka's claim that rationality and physical perfection are essential to human nature, we should look to the practices and judgments of biologists and any others who can plausibly be thought of as having insight into the inner structure of human beings.

The notion that biological kinds are natural kinds certainly has a *prima facie* plausibility and attractiveness. It is probably no coincidence that Aristotle, to whom so much of our traditional thinking about natural kinds and essence can be traced, was also deeply interested in biological classification. Also, Putnam and Kripke themselves often list tigers, cats, and lemons as examples of natural kinds. At least one prominent philosopher of science, Richard Boyd, goes so far as to say that biological species are the paradigm case of natural kinds (1988: 198). Similarly, in her attempt to defend Aristotle's famous but controversial function argument, Jennifer Whiting appeals to the notion that membership in a natural kind is essential to the individual members. Whiting contrasts membership in the class of flutists or prostitutes with membership in the class of men. Individuals are members of non-natural kinds such as flutist or prostitute only accidentally, while they are members of natural kinds essentially.⁴⁸ If human being is a natural kind, then it would follow that individual humans

⁴⁸ See Whiting (35-37, and 46 n12). Although it is unclear whether the Aristotelian concept of man is coextensive with the biological concept of a human, there is at least considerable overlap between the two concepts, and one would expect the Aristotelian concept of man to have a significant biological component. At any rate, what is important to note here is that since man is taken to be a natural kind, one's membership in the class of men is supposed to be an essential property. Since as I shall discuss below there are reasons to doubt that members of biological kinds belong essentially, this understanding of natural kinds will cause problems for the view that biological kinds are natural kinds.

belong to the species essentially.

However, the notion that biological kinds are natural kinds, along with the related notion that we humans belong to the biological kind human being or *homo sapiens* essentially, seems to be based on a naive understanding of systematic biology. If one looks at the actual practice of biologists, one finds an approach to biological classification and explanation that is significantly different from what one finds, for example, in chemistry; thus, whereas the Putnam/Kripke story about natural kinds and essence fits very well the cases of physical-chemical kinds such as gold and water, it does not seem to fit the case of biological kinds. Here is how Elliott Sober describes the current practice of actual biologists and its bearing on essentialism:

To see why essentialism is a mistaken view of biological species, we must examine the practice of systematists themselves. With the exception of pheneticists (whose position will be discussed later), biologists do not think that species are defined in terms of phenotypic or genotypic similarities. Tigers are striped and carnivorous, but a mutant tiger that lacked these traits would still be a tiger. Barring the occurrence of a speciation event, the descendants of tigers are tigers, regardless of the degree to which they resemble their parents. Symmetrically, if we discovered that other planets possess life forms that arose independently of life on earth, those alien organisms would be placed into new species, regardless of how closely they resembled terrestrial forms. Martian tigers would not be tigers, even if they were striped and carnivorous. Similarities and differences among organisms are *evidence* about whether they are conspecific, but a species is not *defined* by a set of traits. In short, biologists treat species as *historical entities*. They do not conceptualize species as natural kinds (Sober 1993: 148).

Although Sober's main target here seems to be views which attempt to classify biological kinds by appeal to morphology the pheneticists whom Sober promises to discuss later turn

out to hold that biological kinds should be grouped according to overall similarity⁴⁹ his comments apply equally to any view which attempts to group biological kinds according to underlying structure, à la Putnam and Kripke. A Putnam-inspired Twin Earth example might help to make the point.⁵⁰ Imagine that the stuff called water on Twin Earth is indeed H₂O instead of XYZ, and that the stuff called gold on Twin Earth has atomic number 79, just like the stuff we call gold. Also assume that human beings on Earth and what are called human beings on Twin Earth have the same underlying structure i.e., the same DNA or chromosome structure. Gold on the two planets and water on the two planets would belong to the same natural kind, since it is molecular structure which determines membership in such kinds. However, the people on Twin Earth and the people on our Earth, even though we are virtually identical in both underlying structure and in more easily observable ways, are nevertheless *not* members of the same species, since the two groups have been, throughout their evolutionary development, causally independent of and isolated from one another.⁵¹

At this point, one might wonder whether Sober's understanding of biological kinds as historical entities is too restrictive. Surely, it might be suggested, there are other ways of understanding the notion of a biological kind such that it is a natural kind. However, in a

⁴⁹ See Sober 1993 (158, 163-64) for further discussion of this approach.

⁵⁰ Putnam famously appeals to the notion of Twin Earth in arguing for his various claims about meaning, reference, necessity, and essentialism. The basic idea is that Twin Earth is just like our Earth in every way, except that the stuff on Twin Earth which looks, smells, feels, and tastes like water does not consist of H₂O; it has some other underlying structure which Putnam refers to as XYZ. The question is, is the stuff on Twin Earth water? Putnam thinks we would say no, and that is why water is essentially H₂O. See Putnam 1975a for an example of Putnam's use of Twin Earth.

⁵¹ For more on Sober's criticisms of the extension of the Putnam/Kripke account of natural kind terms to biological kinds, and for more on the death of essentialism in biology, see Sober 1980 and 1993 (145-149). John Dupré 1981 is also very critical of the attempt to extend the Putnam/Kripke account to biological kinds

recent article Joseph LaPorte argues persuasively that all of the viable contenders for the best account of what makes an individual a member of a biological species undermine essentialism. Specifically, he argues that all of these accounts of speciation entail that individuals are members of biological kinds only accidentally, and not essentially. He lists three basic views: the interbreeding approach, according to which species are groups of interbreeding natural populations that are reproductively isolated from other such groups; the ecological approach, according to which a species is characterized as a lineage with a unique adaptive zone, or ecological niche; and the cladistic approach, according to which an individual species is a lineage of organisms between two speciation events, or between one speciation event and one extinction event (LaPorte: 101-02). The last two accounts require a little explanation. With regard to the ecological approach, a lineage can acquire a unique adaptive zone when it splinters off from a larger group, moves to a new location, and undergoes a change in diet or predators. The cladistic approach is a bit more complicated. The idea here is that speciation occurs whenever a lineage branches off from a larger group and achieves something like reproductive isolation or a unique ecological niche (or some combination of the two). So far, cladism is not easily distinguishable from the other two approaches. The important difference with the cladistic approach is that, when there is a speciation event and a new lineage splinters off from the larger, ancestral group A, not only is there a new splinter species B, but A itself ceases to exist. That is, those left over from group A who do not join the members of B in their new environment, will constitute a new species C.

Now, I think it is fairly easy to see why the cladistic approach suggests that individuals are not essentially members of biological kinds. Members of species C belong to

that species not because of their underlying structure, but because of the activities and behavior of the members of species B. If the members of B had not splintered off and formed a new species, then the members of C would be in A. Thus, membership in a species can be a contingent, accidental affair; in other words, there could be a possible world in which one is a member of a different biological species. A similar conclusion seems to hold for the interbreeding and ecological approaches; for whether a lineage achieves reproductive isolation or a unique ecological niche seems to be a contingent matter. At the very least, it seems that a population can achieve either of these without undergoing an alteration in what could fairly be called underlying structure. Thus, contrary to Putnam's suggestion that biological kinds share an underlying structure,⁵² there seems to be nothing of the sort holding such kinds together. Moreover, even if we insisted on giving these various accounts of speciation an essentialist interpretation, the best we could do would be to say that an individual is, in virtue of being a member of a certain species, capable of interbreeding with others in the species, or that the individual occupies a place in a certain ecological niche. Even if this were an acceptable account of essentialism, it would be extremely difficult to see how we could ever get from this the notion that human beings are essentially rational, as Hurka wants to conclude.

I want to end this discussion of biology, natural kinds, and essentialism by considering a more general argument made by Sober in *Evolution, Population Thinking, and Essentialism*. In this article, Sober gives a general argument for why essentialism cannot find a home in modern evolutionary biology. This brief discussion will serve two purposes: (1) it will offer more evidence of the futility of appealing to biology, in general, to establish a

⁵² See for example Putnam 1970 (141) and 1975a (240).

human essence; (2) it will help to illustrate, in particular, why physical perfection and rationality are not essential to human nature. According to Sober, the most important difference between Aristotelian-inspired, pre-Darwinian biology and modern biology has to do with the appeal to the Natural State Model as a way of explaining biological phenomena.⁵³ According to the Natural State Model, every organism has a natural state to which it aims.⁵⁴ The only thing that can stop an individual organism from reaching its natural state is the presence of an interfering force. An interfering force causes an individual to develop into an imperfect member of the species, or in extreme cases, into a monster. This is how Aristotelian biology explained variation among organisms; without interfering forces, all individuals would naturally develop toward their natural state.

Although the Natural State Model remained the dominant explanatory scheme in biology into the 17th and 18th centuries, and, according to Sober, continues to influence even post-Newtonian mechanics (1980: 353, 360), the model has been rejected in modern biology, and replaced by population thinking. While population thinking is a vast and complicated subject, for our purposes I think we can get by with the following description:

According to the Natural State Model, there is one path of foetal development which counts as the realization of the organism's natural state, while other developmental results are consequences of unnatural interferences. Put slightly differently, for a given genotype, there is a single phenotype

⁵³ See Sober 1980 (360-65) for a more detailed description of the Natural State Model.

⁵⁴ Another problem for Hurka is the fact that Natural State Theory, at least as described by Sober, seems to assume a natural tendency doctrine, according to which an organism has a natural tendency to pursue its good. Recall that Hurka rejects natural tendency doctrines as an undesirable accretion to perfectionism, which the view is better off without. However, without the natural tendency doctrine, it's hard to see how an appeal to biology will allow us to pick out the right sorts of states i.e., physical perfection and rationality.

which it can have that is the natural one. Or, more modestly, the requirement might be that there is some restricted range of phenotypes which count as natural. But when one looks to genetic theory for a conception of the relation between genotype and phenotype, one finds no such distinction between natural state and states which are the results of interference. One finds, instead, the *norm of reaction*, which graphs the different phenotypic results that a genotype can have in different environments (Sober 1980: 374).

The idea here is that, for a group of organisms such as stalks of corn which share a common genotype, all we can do is come up with a graph which measures the various phenotypes which will result in various environments. What we think of as sickly corn is simply, from the point of view of population thinking, that phenotype which results from the combination of a certain genotype with a certain environment (say, one with a certain type of soil, or a lack of rain, or the presence of certain insects or parasites). While the Natural State Model assumes that there is some phenotype which is the natural one independently of a choice of environment, and environments, in turn, are considered natural to the extent that they allow an organism to reach its natural state, population thinking assumes that any environment or any resulting phenotype is just as natural as any other.

It is now clear why we cannot expect modern biology, with its acceptance of population thinking, to account for Hurka's notion of physical perfection. On Hurka's view, someone like Michael Jordan or Charles Atlas exemplifies a great deal of physical perfection, while people such as Stephen Hawking and FDR exemplify considerably less. However, if we take the assumptions of population thinking seriously, all we can say is that the Michael Jordan or the Stephen Hawking that we see is simply the phenotype which results from a certain environment. If the environment were different, then we would have different phenotypes (perhaps Jordan would be confined to a wheelchair, while Hawking would be a

successful tennis pro). However, there is no basis for thinking that one is more *natural* than the other. Although the biologist can perhaps explain why certain environments give rise to certain phenotypes, and perhaps can help us to predict and control which phenotypes get developed via environmental manipulation, the biologist seems unable to give us any special insight into the question of which phenotypes are worth promoting or striving for.⁵⁵

Moreover, this description of modern biology, with its reliance on population thinking, should give us pause whenever we are tempted to appeal to biology to help us to determine the good of humans, or any other species. For it seems that there is little or no room for claims about health or fitness in modern biology. As Sober himself says, our current concepts of function and dysfunction, of disease and health, seem to be based on the kinds of distinctions recommended by the Natural State Model (1980: 377). While Sober speculates that our notions of fitness and health might be understood in some other way (1980: 378), he does not attempt to explain how such an account would go. Also, the reader should recall that, in discussing speciation and essence, Sober told us that if there were creatures virtually identical to us on some other planet (with regard to both genotype and phenotype), these other creatures would count as a different species so long as there had been no causal interaction in the development of the two groups (1993: 148). However, most of us would think that the good of the so-called Twin Earthlings and our own good would have to be virtually identical; the fact that we are of different species seems to cut *no* ice in determining what the good is for members of the two groups. These are the sorts of considerations which ought to make us extremely skeptical of *any* attempt to ground the

⁵⁵ Kitcher makes similar remarks along these lines in *Essence and Perfection*, both with regard to rationality and physical perfection. See section IV of Kitcher's essay for a discussion of physical perfection, and section V for his discussion of rationality.

human good in biological facts or claims.

V. The Substance of Hurka's Conception of the Good

Let's stipulate for the sake of argument that Hurka is right in his claim that human beings are essentially rational. Let's also assume that those who achieve more in the way of physical perfection — such as Charles Atlas and Michael Jordan — also exemplify, in virtue of their physical perfection, their human essence to a greater degree than do those of us who are less physically fit. Given these assumptions, is it plausible to suggest that the human good consists in the development of these characteristics to a high degree? In this section, I shall argue that the claim is implausible and unattractive. The obvious question for Hurka's account of the good is where states such as pain, fear, anxiety, depression, pleasure, happiness, joy, and contentment fit in. This is probably the most obvious, but also the most important, problem for Hurka's (or perhaps *any*) version of perfectionism. I take it that, if we are committed to anything with regard to the human good, we are committed to the notion that some of these states are good for one, while others are bad. Any plausible conception of the good must have at least something to say about such states. A view which offers a *unified* account of the good will have to explain what it is that states such as pleasure and contentment have in common with other non-subjective goods — e.g., rationality and physical fitness.⁵⁶ If we are unable to come up with a plausible unified account, then we will have to

⁵⁶ There are various ways in which one might develop a unified account of disparate goods. One way would be to claim that the various goods involve the development of human nature; this is Hurka's method. Desire-satisfaction theorists achieve unity by claiming that all of the different goods are desired, or would be desired under suitably idealized conditions (thus, I understand desire-satisfaction views to be non-pluralistic in a fairly straightforward sense). Hedonists claim that all of the good things share the property of being types of pleasure, or the property of being instrumentally valuable for their tendency to lead to the

either trim our list of goods or adopt a pluralistic conception of the good. Hurka realizes that this is a concern for his view, and suggests that we might fall back to a pluralistic account of the good.

In order to fully appreciate the way in which Hurka's view is ill-equipped to deal with subjective goods and evils, it will be helpful first to look a little more closely at physical perfection and rationality, so that we can see exactly what they entail. We'll begin with physical perfection. Hurka starts by appealing to the familiar notion that our bodies are made up of various organs and systems, each of which has a function or role to perform (1993: 37-39). The digestive system converts food into nutrients and removes waste, the circulatory system distributes nutrients and oxygen to various parts of the body, the reproductive system allows us to procreate, and so on. If any of these systems fails to serve its purpose, then this detracts from physical perfection. One reaches a high level of physical perfection when all of the various systems—the muscular, circulatory, digestive, reproductive, and nervous systems—function in the way they are supposed to. Moreover, Hurka argues that we reach the height of physical perfection when we achieve great athletic feats, such as running 100 yards in 9.86 seconds or making a long jump of 29 feet (1993: 39). Even though few of us can expect to achieve the athletic prowess of an Olympian, we can all measure our physical perfection by the extent to which we can approximate such feats, and achieve the sorts of functioning with regard to our bodily systems one typically associates with such achievements.

I find Hurka's claim that physical perfection, as thus described, is intrinsically valuable to be attractive and plausible.⁵⁷ However, one is left wondering what to think about

experience of pleasure.

⁵⁷ Hurka thinks that physical perfection is both intrinsically and instrumentally valuable. It's intrinsically valuable because it represents the development of important

a case where someone achieves a good deal of physical perfection but little in the way of pleasure, happiness, or contentment. Now perhaps Hurka would say that such a case is highly improbable, if not impossible; for physical perfection and happiness or contentment are at least highly correlated. If someone is healthy, we generally expect that they will also be fairly happy. Certainly, if one is unhealthy, and one's bodily systems are seriously malfunctioning, then we typically expect this to be accompanied by a considerable amount of pain, anxiety, and unhappiness. Similarly, when people achieve great athletic feats, such as slam-dunking over Shaquille O'Neal or winning the 100 meter race at the Olympics, we typically expect those achievements to be accompanied by a good deal of pleasure, happiness, or a certain sense of satisfaction. Moreover, since people don't typically pursue these types of activities unless they enjoy them at least to some degree, it is quite rare for someone to combine this sort of achievement with an extreme lack of enjoyment. However, this *needn't* be the case. It seems perfectly conceivable that someone might be quite healthy without being particularly happy. Moreover, it seems that great athletes can also suffer from fairly serious cases of depression or anxiety. It could even be the case that it is the drive or desire to achieve great feats that is the major cause of the depression or anxiety. If the pain, depression, and anxiety is severe enough, and the athlete as a matter of fact only experiences a modest amount of pleasure in the achievement of her feat, it may very well be the case that the athlete would be better off, all things considered, if she didn't pursue and ultimately exemplify such physical perfection. At the very least, it seems that one must take such subjective states into consideration when evaluating the value of a life, or the advisability of embarking on a certain

aspects of our humanity. It is instrumentally valuable due to its ability to help us to develop our capacities for rationality—e.g., we may very well think and reason more clearly and more efficiently if we exercise (1993: 39).

course of life.

Of course, physical perfection is only one part of Hurka's conception of the good, so we need to ask what else he might say about the place of subjective goods and evils in his perfectionism. Thus, we need to look more closely at Hurka's conception of rationality and the role that it plays in his view. For Hurka, there are two types of rationality: theoretical rationality and practical rationality. Each is taken to be essential to human nature, and a life in which such rationality is exercised to a high degree is a life that is better, because it is a life more fully human. Here is how Hurka describes the core of the Aristotelian idea that humans are essentially rational:

...Humans are essentially rational because they can form and act on beliefs and intentions. More specifically, they are rational because they can form and act on sophisticated beliefs and intentions, ones whose contents stretch across persons and times and that are arranged in complex hierarchies. These last features distinguish human rationality from that of lower animals. Animals have isolated perceptual beliefs, but only humans can achieve explanatory understanding. They can grasp generalizations that apply across objects and times and can use them to explain diverse phenomena. A similar point holds for practical rationality. Animals have just local aims, but humans can envisage patterns of action that stretch through time or include other agents and can perform particular acts as means to them. By constructing hierarchies of ends, they can engage in intelligent tool use and have complex interactions with others (1993: 39).

The idea here is that we achieve *theoretical* perfection when our instances of knowledge (i.e., our justified true beliefs) are arranged and ordered in a certain way. Similarly, we achieve perfection with regard to *practical* rationality when the goals that we achieve are arranged and ordered in a certain way. If the instances of knowledge (in the case of theoretical rationality) and our achieved goals (in the case of practical rationality) have the right kind of

unity and hierarchical ordering, then we achieve a high degree of perfection.

Hurka measures our rationality along two main dimensions: number and quality. Number is straightforward: we score higher on the dimension of number when we have more justified true beliefs (in the case of theoretical rationality) and more successfully achieved goals (in the case of practical rationality). However, the number of our beliefs and goals is not the only important thing; they also have to be evaluated in terms of quality. Hurka measures the quality of our beliefs and goals by appealing to the concepts of extent and dominance (114-116). He uses these concepts to flesh out a *formal* conception of rationality, which is contrasted with *material* conceptions, according to which rationality requires a certain content. For example, a material conception of theoretical rationality might favor knowledge of certain exalted things, such as God or the heavens. A material conception of practical rationality might require that our goals have a certain moralistic content; for example, they might need to be directed at the common good, or at least be other-regarding to some extent. Hurka thinks that the sort of rationality that can plausibly be thought of as essential to human nature will have to be formal, not material; in other words, it must measure the worth of beliefs and ends using only formal criteria, and none tied to their specific content (1993: 114).

For Hurka, human rationality is measured in terms of sophistication, and sophistication is cashed out in terms of extent and dominance. Another way of putting this is to say that what makes certain beliefs⁵⁸ or goals more important or valuable is that they are

⁵⁸ For the sake of convenience I will talk about beliefs instead of instances of knowledge. However, it is fairly clear that, for us to reach true theoretical rationality, our beliefs have to be instances of knowledge (i.e., *justified, true* beliefs, or whatever the correct account of knowledge entails). Similarly, in talking about practical rationality, I will simply refer to goals instead of achieved goals, but we should keep in mind that genuine

more *general*. For example, Hurka thinks that the knowledge of the number of redheads in Beiseker, Alberta is less valuable and important than knowledge of a fundamental law of the universe because the knowledge of the number of redheads is isolated and particular, while the belief about the universe is more general. It is more general in two senses: it has greater extent, and it has greater dominance. Beliefs have greater extent when their content extends across times and objects; for example, the fundamental law of the universe – assuming there is such a thing – has tremendous extent because it applies to *all* times and objects, whereas the belief about redheads in Beiseker only applies to one particular place and time. Similarly, a goal or intentional state has greater extent if its content stretches across times and objects, including persons. For example, a political leader might intend to adopt a policy which will affect a wide range of persons over a long period of time. Such a goal would have much greater extent than the goal to tie one's shoelaces.

The second measure of generality is hierarchical dominance. The idea here is that a certain belief can be more important or valuable – because of its place in an explanatory hierarchy. Someone who knows it can use it to explain many other truths, which then become subordinate to it in her theory of the world (1993: 115). We can also see how certain practical goals or intentions can have greater hierarchical dominance. For example, the political leader's intentions have greater dominance than those of the shoelace-tie's because, to achieve his end, the political leader must devote a large portion of his life to it, pursuing many other ends as subordinate means to it. Whereas he develops a complex goal-structure with reform at the top and many others below it, the lace-tie's end is accomplished in a few

perfection with regard to practical rationality requires successful pursuit of one's goals. This isn't to say that one cannot achieve a *certain* degree of rational perfection by constructing false, but complex, theories or unsuccessful, but extremely complicated, plans of action.

bodily movements (115-16). Thus, we can see why the lives of philosophers, physicists, and political leaders are often held up as especially good lives; on Hurka's view, they all exemplify a high degree of rationality, and are thus a way to realize and develop one's essential humanity.⁵⁹

My reaction to Hurka's account of rationality and its importance is quite similar to my reaction to his claims about physical perfection. I agree that there seems to be something intrinsically valuable about both theoretical and practical rationality. Moreover, the idea that this value is somehow grounded in its unity and complexity is also attractive. As Hurka points out, this sort of sophistication is a large part, if not the main part, of what sets us apart from the animals. The idea of refusing to develop and exercise these higher capacities seems to be an affront to our sense of human dignity. Like Mill, I suspect that few of us, if given the chance, would give up our higher capacities in exchange for a life of pure animal bliss.

Nevertheless, there is still the question of how subjective goods such as pain and pleasure fit into Hurka's account of the human good. Just as it seems possible that one could achieve high levels of physical perfection without being very happy or contented, it seems that one could have a life which exemplified a great deal of extent and dominance, without bringing much in the way of happiness or contentment. Indeed, it seems especially easy in the case of rationality to imagine such cases. For Hurka's account requires us to put a great

⁵⁹ We can emphasize the formal character of Hurka's conception of rationality by noting that someone like Himmler, who ran the Nazi concentration camps, achieved a great deal of practical perfection, since his plans for murdering the Jews were characterized by a very high degree of extent and hierarchical dominance. For purposes of measuring rationality, it doesn't matter whether the goal is the annihilation of the Jews or the establishment of a national health care system. To the extent that Himmler can be criticized, it is because his plans entailed that others would not achieve their own physical and rational perfection. Thus, Hurka appeals to agent-neutrality to generate familiar other-regarding moral reasons or obligations (1993: 62-68).

premium on complexity, and it would not be at all surprising if the life with the most complexity i.e., with the most extent and dominance were also one of the most anxiety-ridden. Although many people derive great pleasure and satisfaction out of developing, pursuing, and achieving a highly complex network of goals, there are many others who would find such an approach to life to be extremely stressful and anxiety-producing. Just as the pursuit of physical perfection must be tempered by an appreciation of the importance of pleasure, happiness, and contentment, our pursuit and valuing of both practical and theoretical rationality must also be tempered by an appreciation of such subjective goods.

One might think that there is some confusion in thinking that practical rationality can lead to serious unhappiness, since practical rationality is concerned with the setting and achievement of goals, and happiness is a very common goal for human beings to have indeed, happiness might very well come close to being a universal goal. Surely, one might think, Hurka's conception of practical rationality must allow for the centrality, or at least the inclusion, of subjective goods in this way. While it is true that Hurka's formal conception of practical rationality puts no constraints on the content of our goals, and thus allows pleasure and the avoidance of pain a place in our network, it is nevertheless the number and complexity of our goals that is of fundamental importance. Thus, on Hurka's view, the fact that I adopt happiness or pleasure as a goal, or even as my overriding goal, does not guarantee that the happy or pleasant life will be the one where I achieve the most practical perfection. For it could be the case that there is some other network of goals open to me such that pursuit of them would result in a great deal more extent and dominance, without much in the way of pleasure or happiness. In such a case, what really benefits me i.e., what allows me to achieve the most perfection is the less pleasurable, less happy life.

Thinking of practical perfection in terms of competing systems of actions and intentional states can further help to reveal the implausibility of Hurka's conception of the good. Assume that I have two options for the organization of my life, and each option consists of a network of successful goals. Network 1 scores very high along the dimensions of extent and dominance, and also leads to a fair amount of pleasure, happiness, and contentment. Network 2 scores slightly higher than Network 1 along the dimensions of extent and dominance, but also is accompanied by a significant amount of pain, unhappiness, and anxiety. Intuitively, most of us would say that it would make the most sense to go with Network 1, even though we thereby exercise slightly less practical rationality. However, Hurka's view requires that we choose Network 2, since subjective goods and evils such as pleasure and pain are only *instrumentally* valuable or dis-valuable. As implausible as it is to suppose that such states are not intrinsically good or bad, it seems even more implausible to think that a slight increase in extent and dominance could compensate for extreme increases in pain or severe losses in happiness or contentment. Although we might very well agree with Hurka that physical perfection, theoretical rationality, and practical rationality all have a value which is independent of any good feelings they may lead to, or of any desire we might have for them, they are nevertheless goods which have to be balanced against familiar subjective goods and evils.

Hurka himself, although he does not discuss the possibility of lives with high degrees of perfection also being extremely painful and severely unhappy, nevertheless acknowledges that the most perfect life need not be the most pleasant or satisfying life. Indeed, Hurka thinks that the divergence of satisfaction or pleasure from perfection is actually a strength of his view:

Far from damaging perfectionism, this divergence is a source of strength. Imagine a research scientist. If she is dedicated to her work, she will experience frequent frustration as well as sometimes the thrill of successful discovery. Her desire always to be advancing knowledge—a desire not even the best satisfy—may make her life less contented as a whole than if she had simpler wants. Do we then think her wrong to pursue science? Would her life be better with more easily satisfiable desires? Surely not. Her scientific talent is what is best in her and what she should most strive to develop. Or consider the lives in *Brave New World*. They are extremely satisfying, on any plausible account of satisfying, yet lack perfectionist goods such as knowledge and autonomy. If morality had to concern satisfaction, these lives would have to count as ideal, yet we all find them repellant (1993: 27).

Hurka continues in this vein, emphasizing the difficulty of the life of perfection:

If a strong desire or pleasure doctrine [i.e., a doctrine that holds that human beings have a natural tendency to desire perfection above all else, or take their greatest pleasure in perfection] were true, pursuing excellence would be easy. Once we knew where our greatest good lay, achieving it would be just a matter of following our strongest want or enjoying our greatest pleasure. This is not our experience. For most of us, achieving the good requires discipline and concentration. It requires formulating a valuable project and sticking to it despite distractions and temptations. Given this, perfectionism without [natural tendency doctrines] not only matches our moral convictions but also fits our experience of seeking a valuable life (27).

I agree with the sentiments underlying both of these quotations. However, I also believe that if we think of Hurka's research scientist as being, not merely frustrated and unhappy, but severely depressed and anxiety-ridden, then it becomes much less plausible to claim that she should not give up her life of scientific research for something simpler. Moreover, this simpler life does not have to be anything like the lives of the Deltas in *Brave New World*. It just has to be a life which achieves a greater balance between perfection and happiness. Just as the scientist must strike a balance between theoretical perfection and happiness, the Deltas

would be better off with a life which added a significant measure of knowledge and autonomy; however, this is not to say that any accompanying loss of satisfaction and pleasure would not be a genuine loss of something intrinsically valuable for the Deltas.

Hurka is aware that many will find an account of the human good that doesn't include such subjective goods as happiness, satisfaction, pleasure and contentment as intrinsic goods to be implausible and unacceptable. Indeed, Hurka emphasizes that he is not necessarily committed to the view that pleasure or satisfaction is not valuable. He thinks that, if we insist on including these as intrinsic goods, then we will be driven to *pluralism* about the good:

My claim is not that satisfaction has no value. Pure perfectionism makes this claim, but there is also the possibility of a pluralist theory that weighs perfectionist ideas against others about, for example, pleasure or desire-fulfillment. Such a theory can combine these ideas in different ways. It can treat satisfaction as simply another value alongside perfection, or it can say that satisfaction has value only, or has the most value, when it is satisfaction *in* perfection, for example pleasure in scientific research. But it is no objection to a pluralist theory that it does not treat satisfaction as the only value. It may not even be an objection to pure perfectionism that it does not treat satisfaction as any value. (I am uncertain about this.) It is, however, a decisive objection to a pure satisfaction-based morality that it does not treat perfection as a value – this makes the morality unacceptable (1993: 27-28).

Hurka's central purpose in *Perfectionism* is to develop the best and most plausible version of pure perfectionism. Pure perfectionism limits the good to physical perfection and rationality. If we think that our list of goods must be longer and include other goods not reducible to perfection, then we simply adopt pluralism about the good.⁶⁰ However, even if we go this

⁶⁰ In *Virtue, Vice, and Value* Hurka offers a pluralistic account of the good, which includes pleasure, knowledge, and achievement as fundamental goods. It's not entirely clear if Hurka is merely stipulating this list for the sake of argument – that is, for the sake of working out his view about the nature of virtue, which is the central project of that book. It

route, Hurka thinks that we have at least established the proposition that, regardless of what other goods we have to countenance, we must accept perfectionist goods.

Although I am sympathetic with the move toward pluralism about the good, I want to end this chapter by raising some concerns about Hurka's acceptance of pluralism as a fallback position. Recall that Hurka wants to give a unified account of the human good. Instead of simply appealing to something like an objective list view⁶¹ or what Kitcher would call bare objectivism, Hurka looks for some underlying principle or rationale which might unify and explain our various intuitions about the human good and morality. The underlying rationale for Hurka is the notion of human nature, or the human essence. One's good consists in the development of one's essence. Although I have criticized the very notion of a human essence, along with the claim that rationality and physical perfection could plausibly be thought of as part of our essence, in this section I have granted Hurka's claims about the human essence for the sake of argument. Unfortunately, the move toward pluralism threatens to undermine much of Hurka's rationale for the acceptance of perfectionist goods. The admission that there are goods which are not tied to the human essence, along with Hurka's earlier admission that certain essential properties are not plausible candidates for constituents of our good, suggests that the appeal to essence and human nature is doing little or no explanatory work in Hurka's

could be that, in making the case for thinking of virtue as loving the good, Hurka thinks that he needs to make use of uncontroversial, commonsense views about the good, leaving aside the question of whether these are the goods we would accept upon reflection. Another possibility is that Hurka is in fact committed to pluralism, and is merely using *Perfectionism* as an opportunity to push pure perfection as far as it can go. Hurka's parenthetical comment in the quote above suggests that he is genuinely uncertain about the pure-perfectionism vs. pluralism issue.

⁶¹ This terminology is due to Derek Parfit (493-502). I understand this view as holding that there is a list of objective goods (such as friendship, knowledge, virtue, and pleasure), but that there is no unifying principle tying them together.

view.

Early in *Perfectionism*, Hurka discusses a number of essential properties which could not plausibly be thought of as contributing to, or related to, the human good (14-17). Many of these properties are essential to a human being *qua* object or *qua* physical object. For example, *qua* object, I am essentially self-identical. Also, I essentially have the property of being red if red. There are additional properties which are essential to me *qua* physical object, such as being composed of elementary particles and occupying space. Now Hurka rightly points out that it seems extremely implausible to suggest that the development of such essential properties is part of the human good. Thus, in order to avoid what he calls the *wrong properties objection*, Hurka holds that the human good consists in the development of a subset of essential properties; that is, it consists in the development of those properties essential to humans and conditioned on their being living beings (17).

One might think that the restriction of the human good to the development of a subset of essential properties is objectionably *ad hoc*, especially given Hurka's rejection of the possibility of understanding the human good as consisting in the development of those properties that are essential to and *distinctive of* human beings (10-14). However, so long as he can give a complete and unified account of the human good by appeal to the relevant subset of essential properties, then this move is considerably less *ad hoc*, and probably offers enough explanatory power to make it attractive. Unfortunately, we have seen that the view cannot offer a complete and unified account of the human good. It leaves out subjective goods, and seems to lead to value pluralism. Thus, if we accept pluralism, the appeal to essence ends up doing very little explanatory work in Hurka's view, since being an essential property is neither necessary nor sufficient to tie the development of that property to an

individual's good. Hurka seems to be relying very heavily on certain intuitions about what sorts of things are good, and manipulating his conception of the human essence in order to fit those intuitions.⁶²

Moreover, without the unifying power of the notion of human nature or the human essence, it is not at all clear why the skeptic must accept Hurka's claims about perfectionist goods being even *part* of a pluralist morality. Although we might find Hurka's intuitions plausible and attractive on this score, someone who started off being skeptical about perfectionist goods such as rationality and physical perfection could quite legitimately complain that, without the notion of human nature tying all of our judgments about the good together, there is little reason, beyond the bare intuitive notion that rationality and physical fitness are admirable and good, to insist on including rationality and physical perfection on our list of fundamental goods. Thus, the admission of subjective goods seems to undermine much of Hurka's account of the intrinsic value of rationality and physical perfection. Perhaps there are other things we can say about the basis for our intuitions about the value of rationality and physical perfection. I shall argue in a later chapter that there is more we can say about the value of such states, but that we can expect only a very limited degree of explanatory unity in any plausible account of the good. Hurka, at any rate, relies very heavily on the alleged power of perfectionism, and its attendant conception of the human essence, to explain and justify a wide range of our judgments about the good and the right. His view fails in this attempt, and to that extent, support for his version of perfectionism is seriously undermined.

⁶² Of course, this is also one of Kitcher's central criticisms of Hurka in *Essence and Perfection*, though Kitcher focuses exclusively on biological properties.

Conclusion

Hurka sees the human good as consisting in the development and exercise of those properties which are essential to human beings, and he understands the human essence to consist in physical perfection, practical rationality, and theoretical rationality. In defending Hurka against some of Kitcher's criticisms, particularly his suggestion that Hurka attempts to avoid difficult substantive questions about the human good via an appeal to the human essence, I identified the following as what I take to be Hurka's main reasons for thinking that the good consists in the development of whatever properties are essential to human nature. First, the notion that the human good consists in the development of what is essential to human nature is, Hurka thinks, intuitively appealing and plausible. Second, there is the long tradition of perfectionist views in the history of philosophy. Third, Hurka thinks that the human-nature view allows us to explain, justify, and unify a wide range of our intuitions regarding the good, the right, and justice. Also, there is the middle section of Hurka's book, where he offers what Kitcher himself describes as an interesting, subtle, and valuable discussion of questions regarding the measurement, aggregation, and balancing of perfectionist goods. These are all components of Hurka's argument for the substantive moral claim that the good consists in the development of human nature. A proper evaluation of Hurka's position requires that we examine the extent to which these reasons are plausible or persuasive.

However, when we consider these substantive arguments for perfectionism, we see that they fail to establish the truth or even the plausibility of Hurka's view. The central problem for perfectionism is its inability to account for subjective goods and evils, such as pleasure, enjoyment, pain, and suffering. Hurka himself recognizes this as a potential

problem for his view, and suggests that we might be forced to accept pluralism as a fallback position; but even if we do that, Hurka thinks that he has at least demonstrated the intrinsic desirability of perfectionist states such as physical perfection and rationality. Unfortunately for Hurka, the recognition of the intrinsic value of so-called subjective goods undermines what I take to be his strongest reason for favoring perfectionist goods – viz., that the human-nature view allows us to explain, justify, and unify a wide range of our ethical intuitions. The other reasons Hurka gives in support of perfectionism – the long list of perfectionists in the history of philosophy and the intuitive attractiveness of the idea of connecting the human good to the development of human nature – constitute quite weak support for Hurka's position. For the idea that the human good consists in the development of human nature ceases to be attractive once it is recognized that such a view leaves no room for subjective goods, while the fact of the popularity of perfectionism throughout the history of philosophy is counteracted by the fact of its unpopularity throughout most of the twentieth century, along with the similarly long list of figures throughout the history of philosophy who have rejected the view. Finally, while the middle part of Hurka's book offers an account of how various perfectionist goods are to be weighed against one another, it offers no account of how perfectionist goods are to be weighed against subjective goods.

Hurka makes two distinct claims. First, he argues that the human good consists in the development of whatever properties are essential to human nature. Second, he claims that physical fitness along with practical and theoretical rationality are essential to human nature. However, it is clear that such properties or characteristics are not essential to human nature, at least if we understand Hurka to mean by 'human being' the biological type *homo sapiens*, and he gives every indication that this is indeed what he means. Furthermore, even if Hurka

were right that there is indeed a human essence, and that his favored properties were part of that essence, his version of perfectionism would offer an impoverished and implausible account of the human good. Of course, this raises the question of whether we might understand human being in some way other than in biological terms, and whether such a conception might allow for a richer array of goods. Specifically, a perfectionist might claim that our good consists in the development of those properties which are essential to us *qua* persons or rational agents. In the next chapter, I shall consider such an account.

Chapter 4 Self-Realization Accounts of the Good

I. Introduction

In this chapter, I shall consider an historically important, but recently neglected, version of perfectionism known as self-realization. Although self-realization versions of perfectionism have their roots in a wide range of figures from the history of philosophy such as Plato, Aristotle, Butler, Reid, and Kant, they are most closely associated with the post-Hegelian British idealists F.H. Bradley and T.H. Green. Bradley and Green both hold that the good or morality⁶³ consists in the perfection or realization of the self. Self-realization accounts are of interest to us here for several reasons. Most importantly, self-realization offers an interesting alternative to Thomas Hurka's version of perfectionism considered in the previous chapter. Whereas Hurka offered a metaphysically stripped-down version of perfectionism, relying only on a purely descriptive, biological conception of human nature and its development, along with the popular and familiar essentialism of Putnam and Kripke, Bradley and Green assume a normative and metaphysically more robust understanding of the

⁶³ This obviously raises questions about the relation between the good of the agent and moral goodness. As we've seen in previous chapters, it is customary to distinguish between an agent's good and the moral goodness of her actions. Utilitarians, for example, typically take the good for the agent to be something like pleasure or desire-satisfaction; however, moral goodness, or preferably moral *rightness*, is a matter of promoting the good of everyone. Thus, when we say that, for Bradley and Green, the good or morality consists in self-realization, we could mean one of two things: (1) that the agent's good consists in self-realization; or (2) that morality, including familiar other-regarding moral demands, consists in self-realization. Unfortunately, in discussing the views of Bradley and Green it can be difficult to avoid this ambiguity, since a large part of the point of self-realization accounts seems to be the blurring of the border between the agent's good and the common good or morality. A clear teasing out of the complex relationship between the good of the agent and morality is one of the burdens of this chapter. But, for now, I'll simply note the ambiguity inherent in the identification of self-realization with the good *or* morality.

self and its realization or perfection. Given the apparent failure of Hurka's stripped-down model to deliver his (in my view) attractive conclusions about the value of physical perfection and rationality, we have good reason to explore the Bradley-and-Green alternative. Another reason to explore self-realization accounts is that, in identifying morality with self-realization, Bradley and Green offer at least a *prima facie* answer to the question, "Why be moral?" The contrast with Hurka is again helpful. Hurka generates other-regarding duty by appeal to agent-neutrality. Bradley and Green, however, in linking morality with self-realization, are more clearly in the eudaimonistic tradition most closely associated with Aristotle, and to that extent, offer a more traditional and purer form of perfectionism.

Another attractive feature of the perfectionism of Bradley and Green is its promise to offer a somewhat richer array of goods than does Hurka's version. Bradley, in particular, explicitly offers a diverse list of goods. For Bradley, self-realization is found in the following: the fulfillment of those duties associated with one's station in society; the fulfillment of certain other duties falling under a category he calls "ideal morality";⁶⁴ the pursuit of truth or knowledge through intellectual inquiry; and the creation of beauty through artistic activity. Also, pleasure gets included as a good, though Bradley is adamant that pleasure is not *the* good. Although Green is not as explicit in granting such a varied list of goods, he nevertheless gives some indication of recognizing a more diverse list of goods than Hurka does. For example, while Green may not think that pleasure is an intrinsic good, he does nevertheless think that there is an intimate connection between self-realization and pleasure, holding that "since there is pleasure in all realization of capacity, the life in which

⁶⁴ It's a bit murky precisely what Bradley thinks ideal morality involves, but he is at least clear that it includes duties not covered by the account offered in "My Station and its Duties." I'll have more to say about the content of ideal morality below.

human capacities should be fully realized would necessarily be a pleasant life.⁶⁵ Thus, although Green may, in the end, resist the notion that pleasure is an *intrinsic* good, his linking of pleasure with self-realization at least takes the edge off his perfectionism for those who find the rejection of pleasure and related states as intrinsic goods unattractive. Moreover, on one reading of Green, self-realization involves the exercise of one's rational capacities in such a way that the agent is able to confer value onto objects of reflective choice. Since it seems that the objects of reflective choice could form a rather varied list, this seems to be a way in which Green's version of self-realization can offer a wide and diverse range of goods.

While the self-realization accounts offered by Bradley and Green have their attractions, there are nevertheless some serious problems with these views. One problem has to do with the metaphysics and epistemology underlying these accounts. As we saw in the last chapter, Hurka's metaphysical commitments, though fairly uncontroversial, were not sufficient to underwrite his conclusions about human nature and the good. Bradley and Green, by contrast, do seem to have metaphysical and epistemological views robust enough to support their accounts of self-realization and the good; however, many of these metaphysical and epistemological views are extremely controversial, and to the extent that Bradley and Green rely on them, their arguments are unconvincing.

A second problem has to do with the diversity of the goods recognized by Bradley and Green. While this diversity is in some ways a decided advantage of self-realization views, it also causes problems. The challenge for Bradley is to give an account of self-realization which can plausibly be thought to underlie and unify the various goods he recognizes. Green's view, which identifies the good with the development and exercise of

⁶⁵ Green §361.

one's capacities for rational deliberation, faces similar problems. There are various ways in which Green's view might be interpreted, and none of them is plausible. On certain ways of understanding Green's account, there are very few, if any, substantive constraints on the content of our conception of the good. These versions have the advantage of offering a potentially diverse list of goods, but they also have the disadvantage of offering little in the way of guidance in developing an acceptable conception of the good, and seem to make the choice of goods arbitrary. Alternatively, we might pursue a more robust, substantive interpretation of Green's view, but such views either seem to presuppose some other account of the good which is independent of the notion of self-realization or they leave out important non-perfectionist goods.

A third problem for self-realization views has to do with the relation between selfishness and self-sacrifice. Although the notion of grounding the demands of other-regarding virtue or morality ultimately in the good of the agent is *prima facie* attractive, it leaves us with a profoundly distorted conception of the demands of prudence and morality. A major upshot of this discussion of self-realization accounts of the good and morality will be a recognition of the attractiveness and plausibility of pluralism both about the good and practical reason.

II. Bradley on Self-Realization

Peter Hylton describes Bradley as "the most influential of the British Idealists" and perhaps the most prominent philosopher in Britain throughout the 1890's (Hylton: 44). His collection of essays, *Ethical Studies*, was the first major statement of the ethical thought of

the British idealists,⁶⁶ and although reactions to Bradley's essays were mixed,⁶⁷ the book remains an important statement of self-realization accounts of morality. Peter Nicholson calls it the best introduction to the ethical thought of the British idealists (Nicholson: 6). One can hardly deny, I think, that the book has its own peculiar virtues and vices. Bradley's essays enjoy the virtue of having been expressed in a writing style which is lively, engaging, and at times quite entertaining; however, many have been put off by Bradley's polemical and, at times, dogmatic mode of argumentation. An advantage, from our perspective at least, of Bradley's presentation of self-realization is that he makes at least some attempt to support his ethical view without appealing to the metaphysics of idealism, often relying instead on a demonstration of what he takes to be the shortcomings and inconsistencies of rival views, while extracting from them some element of truth which can be incorporated into a more plausible account. However, Bradley's presentation has the vice, or at least disadvantage, of proceeding in the Hegelian, Dialectical manner of thesis, antithesis, and synthesis. This Hegelian methodology gives rise to two problems: first, there is a temptation for Bradley to distort views under consideration so that they might more easily fit the Dialectical story he wants to tell;⁶⁸ second, the reader is often tempted to take merely provisional views as

⁶⁶ Green's *Prolegomena* was published posthumously in 1883, seven years after the publication of *Ethical Studies*.

⁶⁷ See, for example, Sidgwick's rather harsh review of *Ethical Studies* in the journal *Mind*. Bernard Bosanquet, by contrast, was enthusiastic in his praise of the work (Nicholson: 51-52).

⁶⁸ In making the case for the relative superiority of Green over Bradley, Brink argues that Green's rejection of the Hegelian Dialectical methodology allows him to make better use of figures and positions from the history of philosophy: Freed of the need to shoehorn other philosophers into a Dialectical moment, Green is able to treat other traditions and philosophers more sympathetically and constructively (Brink 2003: 112).

Bradley's considered position.⁶⁹

Aside from these stylistic and methodological issues, when we examine the substance of Bradley's account we again see a mixture of virtues and vices. Some of these are common to both Bradley and Green, while others are peculiar to Bradley. As I've already noted, one of the main virtues of Bradley's position is that he offers a fairly rich and varied list of goods. Bradley is also successful in capturing many of our considered judgments about what is good and right. For example, the view expressed in *My Station and its Duties* captures the intuition that many of us have that some moral considerations, unlike the forward-looking considerations favored by consequentialism or the backward-looking considerations of familiar versions of deontology, are instead sideways-looking; you find yourself with an obligation to act in a certain way simply in virtue of occupying a certain position in society. Also, because morality is a matter of self-realization, Bradley offers an answer to the question, *Why be moral?* However, there are serious problems with the view as well. For example, in admitting a fairly diverse list of goods, it becomes less clear that these goods are in any real sense unified as aspects of self-realization; in other words, Bradley ends up with a view which seems more like pluralism than self-realization. Another problem with Bradley's view is the degree to which his conception of the self is social in nature; I'll argue that many of Bradley's claims about the ways in which the essence of the self is bound up with its social relations are implausible. Also, Bradley's account of selfishness and self-sacrifice is highly

⁶⁹ The most common mistake along these lines has been for commentators to take the view discussed in *My Station and its Duties* as Bradley's considered view. This is surprising given that Bradley makes it quite clear near the end of that essay (pp. 202-06) that, although he takes the view to be decidedly superior to the sorts of hedonism and Kantianism discussed in earlier essays, it is nevertheless an incomplete and ultimately indefensible account of morality. See Nicholson (6, 39-40) for a discussion of this common misreading of Bradley. See also Wollheim (233-34).

problematic. Although his identifying other-regarding duty with self-realization is attractive insofar as it explains why we have reason to exercise such virtue, it threatens to undermine the commonsense distinction between self-sacrifice and selfishness. A general problem with Bradley's position, which will keep popping up, is Bradley's tendency, notwithstanding his claims not to have a metaphysical system at hand with which to prove his moral conclusions, nevertheless to appeal to controversial idealistic metaphysical and epistemological claims at crucial stages in his argument.

It would perhaps be helpful at this point to briefly lay out Bradley's argumentative strategy in *Ethical Studies*. There are a total of seven essays, the first of which, "The Vulgar Notion of Responsibility in Connexion with the Theories of Free-Will and Necessity," deals with questions of freedom, determinism, and moral responsibility. It is in the second essay, "Why Be Moral?" that Bradley first introduces what he calls the "formula of self-realization" (80),⁷⁰ which holds that morality, or the final end, consists in the realization of the self. Essays III and IV are best thought of as examinations of possible, but ultimately misguided, interpretations of the idea of self-realization: Essay III, "Pleasure for Pleasure's Sake" rejects hedonism as a plausible account of self-realization, whereas Essay IV, "Duty for Duty's Sake," considers and rejects a Kantian-inspired account of self-realization. It is in Essay V, "My Station and its Duties," that we start to get Bradley's positive account of self-realization, and this account is supplemented by Essay VI, "Ideal Morality." The final essay, "Selfishness and Self-Sacrifice," explores some of the implications of Bradley's account for our thinking about morality and self-interest. Since I am mainly interested in Bradley's positive account of

⁷⁰ In this section, parenthetical page numbers refer to Bradley's *Ethical Studies*, unless otherwise indicated.

self-realization, my discussion here will focus on Essays II, and V-VII.

I have said that one of the advantages of self-realization accounts of morality is that they offer an answer to the question, Why be moral? However, Bradley would resist putting the issue in quite those terms. For Bradley, the question, Why be moral? is, strictly speaking, nonsensical or meaningless (64). Bradley argues that there is an illicit, dogmatic assumption behind the question, and that assumption is that morality or virtue is to be chosen for the sake of something else, such as pleasure (60-61). When I ask why I should be moral, I must be assuming that there is some other end which is choice-worthy in itself, and that morality or virtue has only instrumental value toward that end. Bradley argues that all systems of morality must assume some end which is not a mere means. Ethical hedonism, for example, assumes that the ultimate end is pleasure, and justifies virtue insofar as it leads to the pleasure of the agent, in egoistic forms of hedonism, or to the pleasure of everyone, in universalistic versions. Bradley thinks that pleasure cannot be the ultimate end. First of all, the idea that pleasure is the ultimate end and that virtue is choice-worthy only insofar as it leads to pleasure is in direct antagonism to the voice of the moral consciousness (61). Thus, Bradley appeals to common moral opinion in arguing that virtue is valuable as an end in itself.⁷¹ Second, in the essay *Pleasure for Pleasure's Sake* Bradley argues at length that the end presumed by hedonism, the highest sum of pleasures, is a meaningless notion and an

⁷¹ In his critical discussion of a Kantian-like deontological position in the essay, *Duty for Duty's Sake* Bradley continues to make the case that virtue, or the moral good, is an end in itself (142-43). Bradley's main criticism of this pseudo-Kantianism (he makes it clear in footnote 2 on pages 142-43 that he does not intend the view under consideration to be an accurate or complete statement of Kant's actual position) is that the notion of the Good Will as an end in itself is too abstract and lacking in content to be a useful and compelling account of morality or the good. The main positive lesson we learn from a consideration of this pseudo-Kantianism is that morality or virtue must be an end in itself.

illusionary goal.⁷²

Having established (at least to his own satisfaction) that morality assumes some sort of end in itself, and that pleasure cannot plausibly be taken to be that end, Bradley then turns to the specification of the ultimate end. It is in the context of this discussion that Bradley first uses the phrase, self-realization. As Bradley puts it: What remains is to point out the most general expression for the end in itself, the ultimate practical why ; and that we find is the word *self-realization* (64). But what precisely is self-realization, and why should we think that it is the ultimate end? Here, Bradley strikes a somewhat skeptical note:

How can it be proved that self-realization is the end? There is only one way to do that. This is to know what we mean, when we say self , and real , and realize , and end ; and to know that is to have something like a system of metaphysic, and to say it would be to exhibit that system. Instead of remarking, then, that we lack the space to develop our views, let us frankly confess that, properly speaking, we have no such views to develop, and therefore we can not *prove* our thesis. All that we can do is partially to explain it, and to try to render it plausible. It is a formula, which our succeeding Essays will in some way fill up, and which here we shall attempt to recommend to the reader beforehand (65).

This statement of methodology is important for at least a couple of reasons. First, Bradley alerts us to the fact that he intends to appeal to substantive moral argument in order to support his claim that the good is self-realization. He eschews a strategy of first laying out a full metaphysical system and then demonstrating that the system entails that the ultimate end is self-realization. Instead, Bradley will describe his conception of self-realization and explain how it relates to common moral categories, such as goodness, virtue, rightness, obligation, selfishness, and so on. This is a helpful methodological approach, since it allows us to

⁷² The arguments that Bradley offers on this score are very similar to those Green makes against hedonism.

critically evaluate Bradley's claims about goodness and morality without going through the arduous task of evaluating the metaphysics and epistemology of Bradley's idealism. The second important point about this passage is that it gives us a fairly concrete idea of where to look for Bradley's conception of self-realization: we are to look to the later essays in *Ethical Studies*. It is in "My Station and its Duties" and "Ideal Morality" in particular that Bradley lays out his positive account of self-realization.

Before we turn to those essays, we should give a little more consideration to some of the general remarks Bradley makes about self-realization in "Why Should I Be Moral?" In this essay Bradley describes himself as putting forward the formula of self-realization, and hopes that the succeeding essays will offer something like a commentary and justification of the concept (80-81). As we've seen, Bradley argues that any ethical system must assume an ultimate end, something that is desired or pursued for itself. Why assume that the end is self-realization? Here Bradley appeals to a version of psychological egoism:

Let us then dismiss the moral consciousness, and not trouble ourselves about what we think we ought to do; let us try to show that what we do do, is, perfectly or imperfectly, to realize ourselves, and that we can not possibly do anything else; that all we can realize is (accident apart) our ends, or the objects we desire, and that all we can desire is, in a word, self (66).

This claim that only the self can be desired is restated several times in this essay. After explicitly declining to offer a full-fledged account of what it means to desire something, Bradley insists that we say with confidence that, in desire, what is desired must in all cases be self and that we think it is clear that nothing moves unless it be desired, and what is desired is ourself (67-68). After giving a brief sketch of the mechanism by which we come

to associate ourselves with other objects and thereby come to desire ourselves in them,⁷³ Bradley appeals to common judgments regarding desire: By passing by the above [i.e., the brief sketch of the associational mechanism of desire], which we can not here expound and which we lay no stress on, we think that the reader will probably go with us so far as this, that in desire what we want, so far as we want it, is ourselves in some form, or in some state of ourselves; and that our wanting anything else would be psychologically inexplicable (68).

These comments suggest that Bradley is committed to *some* version of psychological egoism.⁷⁴ But what sort of psychological egoist is he, exactly? It is clear that Bradley rejects psychological hedonism, according to which the only thing anyone desires as an end in itself is pleasure (92-93). Unfortunately, beyond that, Bradley's views on egoism are fairly obscure. For example, it's difficult to know how to take Bradley's claim that all that we can desire is, in a word, self (66). What does it mean, exactly, to desire self? Although I have a tolerably clear intuitive notion of what it means to desire something *for myself*, or to desire that which I take to be *in my self interest*, the claim that I *desire self* seems to be, at best, a metaphorical or otherwise imprecise way of saying something else.

Clearly, we need to make better sense of Bradley's egoistic language, and there are at least a couple of possible interpretations of Bradley's claims which suggest themselves. First, he might be claiming that whatever I desire is bound up with myself in the relevant way, in virtue of the fact that I desire it; on this view, self-realization has no meaning independently of the notion of desire-satisfaction. A second way of reading Bradley is to take him as

⁷³ I shall have more to say about this mechanism below.

⁷⁴ C.D. Broad cites both Bradley and Green as examples of psychological egoists who are not psychological hedonists (99). As we shall see, however, it is not clear in what sense, if any, Bradley and Green were truly psychological egoists. Nevertheless, in the above quotations, Bradley certainly *sounds* like a psychological egoist.

claiming that we do have a reasonably clear and coherent notion of self-realization which is independent of the notion of desire-satisfaction, and that it is just a contingent psychological fact about us that the only thing we desire for itself is self-realization, so construed.

Unfortunately, neither of these interpretations will withstand much scrutiny.

The first view, which effectively reduces self-realization to desire-satisfaction, threatens to strip the concept of self-realization of all content if we combine it with the apparently modest assumption that we always do what we want, all things considered.⁷⁵ If I always do what I most desire, all things considered and given the circumstances, and if self-realization is simply a matter of satisfying my desires, then no matter what I do, I will achieve

⁷⁵ The assumption here is that we always do what we most desire in the sense of always acting on our strongest desire at the moment of acting. There is some evidence that Bradley holds such a view: he suggests that, since all action is grounded in desire, one must do what one wants to do (252-53), and he says that the proposition that we only do that which we want is a tautology (255). However, as we shall see in the next section, this is a view which Green clearly rejects: Green thinks that, as rational agents, we can always distinguish between the *intensity* and the *authority* of our desires, and act according to the less intense desire if we so choose. But regardless of the actual views of Bradley and Green, I take it that there is a way of understanding this assumption such that it is compatible with the possibility of weakness of will. Weakness of will occurs when one's desire for, say, a short-term benefit overwhelms one's desire to do what one has judged to be the thing that one ought to do, all things considered. For example, I form the judgment that I ought to lose weight and realize that I can do so only by eliminating sweets from my diet. However, I still have the desire to eat Oreo cookies, and if the desire for the cookies is strong enough, and my desire to act according to my all-things-considered judgment about the importance of losing weight, and hence abstaining from the Oreos, is weak enough, then I will succumb to my desire for the Oreos and thereby exemplify weakness of will. In such a case, I take it that there is a clear and straightforward sense in which I nevertheless do what I most want to do. For the agent who does not suffer from weakness of will, the desire to act according to her judgments about what she ought to do are strong enough that they outweigh or overcome competing desires. The present assumption is also meant to apply to cases of coercion. Even in such cases, there is a sense in which one does what one most wants to do, given the circumstances. If a robber says, 'Your money or your life,' you always have the opportunity to opt for the bullet in the head. Although there is a sense in which you don't want to hand over your wallet, *given the circumstance and the relevant alternatives*, what you most desire is to hand over the wallet. If you didn't have such an overriding desire, then you would not act accordingly.

self-realization. Bradley himself seems to suggest such a conclusion when he says about self-realization that we cannot possibly do anything else (66). On this view, the call to self-realization does not point in any particular direction, and therefore, the notion of self-realization cannot act as any sort of standard or guide to conduct. Not only is this an intuitively unattractive view about self-realization, but it is also in considerable tension with what Bradley later says about the bad self, which is that part of us which desires things other than the good or self-realization (279-80). Whatever we ultimately make of Bradley's claims about the bad self, it is at least clear that he thinks that it is possible for someone to have a desire the satisfaction of which does not lead to self-realization.

Bradley's notion of the bad self also undermines our second possible interpretation of his egoistic claims. According to this second interpretation, we have a reasonably clear and plausible notion of self-realization which is independent of desire-satisfaction, and it turns out that self-realization is what we all desire. Not only does this seem like a patently false empirical claim about human psychology and behavior, but it is also clearly undermined by Bradley's claims about the bad self, which desires things other than the good (and hence, self-realization), and which Bradley tells us is characterized by pride, hate, revenge, passionateness, sulkiness, malice, meanness, cowardice, and recklessness (280). Moreover, if Bradley were serious about the claim that we all necessarily desire self-realization, then it would be unclear why he feels the need to go to such pains to convince us that self-realization, as he understands it, is what we ought to pursue.

Neither of our proposed interpretations of Bradley's claim that all that we can desire is self is convincing. Bradley must be committed to something more subtle and complex than what we have so far considered. In order to fully appreciate Bradley's claims about our

desire for self we will have to look more closely at what he says about the good and bad selves, and at his account of our psychological, and in particular our moral, development. As we shall see, central to Bradley's account of the morality of self-realization is his claim that the bad self i.e., that part of us which desires things other than the good and self-realization as Bradley understands it is not the *true* self, and that the true self is that part of us which desires the good i.e., those things which Bradley associates with self-realization. We'll also see that Bradley's account of our psychological development is central to his claims about the content of the true self.

Let's start with Bradley's views regarding the nature of the self and its development. We should first note what Bradley thinks the self is *not*. Bradley rejects what he takes to be the empiricist account of the self, according to which it is a mere collection of experiences or feelings.⁷⁶ He also rejects the notion of the self as an abstract subject of experience; as we shall see, for Bradley the self essentially consists of a certain socialized content.⁷⁷ Beyond this, Bradley's pronouncements about the self are fairly obscure. He tells us that, in desiring the self, we necessarily desire the self as a whole (68-73). He goes on to describe the self which is the end of self-realization as an infinite whole (74). After offering an argument for the proposition that mind (or the self) is infinite (75-78), Bradley admits that there is nevertheless a sense in which our selves are finite: We admit the full force of the objection. *I am finite; I am both infinite and finite, and that is why my moral life is a perpetual progress. I must progress, because I have an other which is to be, and yet never quite is myself; and so,*

⁷⁶ See for example p. 68. See also *Pleasure for Pleasure's Sake*, where Bradley ties the empiricist account of the self to hedonism.

⁷⁷ This comes out probably most clearly in *My Station and its Duties*. But see also *Duty for Duty's Sake*.

as I am, am in a state of contradiction (78).

Although Bradley has denied having a system of metaphysics at hand which he can use to prove his conclusions about self-realization, when it comes to fleshing out his notion of the self he nevertheless appeals to apparently Hegelian-inspired claims about the nature of persons and their relation to the self as an infinite whole. In response to Goethe's dictum,

Be a whole *or* join a whole, Bradley says, You can not be a whole *unless* you join a whole (79). Bradley continues,

The difficulty is: being limited and so not a whole, how extend myself so as to be a whole? The answer is, be a member in a whole. Here your private self, your finitude, ceases as such to exist; it becomes the function of an organism. You must be, not a mere piece of, but a member in, a whole; and as this must know and will yourself (79)

Bradley wants to make it clear that this coming to be a member of an infinite whole does not entail the complete obliteration of a sense of the separateness of individuals:

...in the moral organism the members are aware of themselves, and aware of themselves as members...I know myself as a member; that means I am aware of my own function; but it means also that I am aware of the whole specifying itself in me. The will of the whole knowingly wills itself in me; the will of the whole is the will of the members, and so, in willing my own function, I do know that the others will themselves in me (80).

These passages are important for several reasons. They illustrate Bradley's reliance on idealistic metaphysical principles to support his claims about morality. They also give us our first glimpse at Bradley's central notion of an individual's *function* within a group or institution. Also, they offer an illustration of Bradley's penchant for combining somewhat exotic metaphysical claims with commonsense views about ethics. In this case, we can clearly see Bradley's wish to combine his conception of the relation between the individual

and society as a whole, with its attendant promise of a harmony of interests, with the everyday, commonsense conviction that individuals are distinct in certain morally important ways. In other words, Bradley's commitment to commonsense moral opinion or judgment forces him to take seriously, at least to some extent, the separateness of persons.

But what is the precise relation between the individual self and the infinite self?

Friends of commonsense moral opinion will not be encouraged by the following:

No doubt the distinction between separate selves remains, but the point is this. In morality the existence of my mere private self, as such, is something which ought not to be, and which, so far as I am moral, has already ceased. I am morally realized, not until my personal self has utterly ceased to be my exclusive self, is no more a will which is outside others wills, but finds in the world of others nothing but self

Realize yourself as an infinite whole means, Realize yourself as the self-conscious member of an infinite whole, by realizing that whole in yourself. When that whole is truly infinite, and when your personal will is wholly made one with it, then you also have reached the extreme of homogeneity and specification in one, and have attained a perfect self-realization (80).

Although Bradley feels compelled to admit that we will continue to have an awareness of our own individuality or separateness, morality requires that we identify our own good with the good of the whole, or at least with the good of others. This account of the proper relation between the individual and the whole is problematic. While it is *possible* that there could be a utopia where the desires, aspirations, and interests of all individuals are in absolute harmony, so that there would never be a conflict between my own good or self-realization and the good or will of others, this seems rather unlikely. At the very least, it is a very great distance from any actual state of human existence, and therefore seems decidedly inappropriate as a standard of ethical conduct for creatures like ourselves. Given the

implausibility and apparent extreme demandingness of this view, it is not surprising that Bradley resorts to controversial metaphysical claims to support it.

Now that we have a better idea of Bradley's conception of the self and I have raised some initial objections to his view, let's take a closer look at the *content* of Bradley's conception of self-realization. In other words, I want to examine what self-realization consists in for Bradley. To do this, we need to consider Essays V and VI, *My Station and its Duties* and *Ideal Morality*. *My Station and its Duties* is interesting in the present context because, even though it identifies self-realization with the fulfilling of one's function in society, at times Bradley makes it sound like this function is typically not very demanding; thus Bradley offers a possible response to a worry raised in the last paragraph. However, *Ideal Morality* suggests a further set of duties which threaten to restore the extreme demandingness of Bradley's view. Thus, once again we will see a basic tension in Bradley's position: although he wants to take seriously the modest demands of *My Station*,⁷⁸ he nevertheless is driven to posit a more severe set of moral demands, and it is not at all clear how the two views are to be reconciled.

To understand the view of *My Station*, we need to be clear on Bradley's account of our psychological and moral development. Although Bradley rejects psychological hedonism, he nevertheless thinks that our desire for pleasure and aversion to pain play a central role in the process by which we come to desire things other than our own pleasure. His basic view seems to be something like the following.⁷⁹ As infants, all that we desire is our own pleasure

⁷⁸ From here on, this is typically how I shall refer to the view expressed in *My Station and its Duties*.

⁷⁹ See Bradley 67-68, 168-172, and especially 280-93 for his description of this process.

and the absence of pain; Bradley seems to think of this desire for pleasure and aversion to pain as an instance of our desire for self. Over time, we come to associate ourselves with certain external objects – these objects can range from such things as a toy or stuffed animal to a nurse or a parent. Because we associate ourselves with these objects, their absence or destruction causes us pain, and their presence or flourishing causes us pleasure. When the child loses track of his teddy bear or mother, he feels that a part of him is missing. Similarly, he experiences an attack on those things close to him as an attack on himself. These feelings continue into adulthood, so that if I discover that a tree I played in or a house I lived in as a child has been destroyed, I feel that a part of myself has been lost. It is through this process of association that we come to, in some sense, identify ourselves with objects which are initially other than ourselves; according to Bradley, this is how we come to have interests in things other than our own pleasure and freedom from pain – i.e., this is how we come to take the promotion, presence, or flourishing of these things to be an end in itself. This is how self-realization can be a matter of the promotion or protection of things or people which, pre-theoretically, we would think of as distinct from ourselves.

Bradley uses this general account of our psychological development to explain how the fulfillment of the duties associated with my station in society can be constitutive of my self-realization. Bradley rejects what he calls individualism, according to which individual human beings have an existence which is independent of the societies of which they are a part, along with the related doctrine that societies, states, and institutions have no reality apart from their status as collections of such individuals (165-71). According to Bradley, the notion of a man abstracted from the various relations which constitute the society to which he belongs is a mere theoretical fiction: ...what we call an individual man is what he is because

of and by virtue of community (166). Bradley continues:

What we mean to say is, that [the individual man] is what he is because he is a born and educated social being, and a member of an individual social organism; that if you make abstraction of all this, which is the same in him and in others, what you have left is not an Englishman, nor a man, but some I know not what residuum, which never existed by itself, and does not so exist. If we suppose the world of relations, in which he was born and bred, never to have been, then we suppose the very essence of him not to be; if we take that away, we have taken him away; and hence he now is not an individual, in the sense of owing nothing to the sphere of relations in which he finds himself, but does contain those relations within himself as belonging to his very being; he is what he is, in brief, so far as he is what others also are (166-67).

Bradley is not just claiming that the content of my individual personality is largely a function of the environment in which I am raised. This would be a comparatively modest claim, and few would be willing to deny it. Bradley is making the more radical claim that the various relations which tie me to my friends, family, acquaintances, community, state, and nation are an essential part of my being.⁸⁰

Bradley's account of our psychological development helps to explain how it is that these various institutional relations come to be part of one's very being. Although Bradley thinks that some of the traits which constitute one's being are inherited from one's parents, he seems to think that the influence of environment and education are the primary determinates. The influence of one's parents and others in society is so strong during infancy that there is no way a child could not take on those traits which make him essentially a member of a certain group. As Bradley puts it, the tender care that receives and guides him is impressing on him habits, habits, alas, not particular to himself, and the icy chains of universal custom

⁸⁰ See also pp. 168, 172, 173.

are hardening themselves round his cradled life (171). By the time a child is old enough to conceive of himself as a separate individual, his self, the object of his self-consciousness, is penetrated, infected, characterized by the existence of others (172).

Bradley is quite clear that he conceives of the identity of the self in rather particular terms. An Englishman is essentially an Englishman, a German is essentially a German, and so on. About the self so conceived, Bradley says the following:

Its content implies in every fibre relations of community. He learns, or already perhaps has learnt, to speak, and here he appropriates the common heritage of his race, the tongue that he makes his own is his country's language, it is (or it should be) the same that others speak, and it carries into his mind the ideas and sentiments of the race (over this I need not stay), and stamps them in indelibly. He grows up in an atmosphere of example and general custom, his life widens out from one little world to other and higher worlds, and he apprehends through successive stations the whole in which he lives, and in which he has lived. Is he now to try and develop his individuality, his self which is not the same as other selves? Where is it? What is it? Where can he find it? The soul within him is saturated, is filled, is qualified by, it has assimilated, has got its substance, has built itself up from, it *is* one and the same life with the universal life, and if he turns against this he turns against himself; if he thrusts it from him, he tears his own vitals; if he attacks it, he sets his weapon against his own heart. He has found his life in the life of the whole, he lives that in himself, he is a pulse-beat of the whole system, and himself the whole system (172)

Since the self is constituted by the relations which tie it to its society, the self is realized through fulfilling those functions associated with one's station in the social organism. As Bradley puts it, an individual man is one of a people, he was born in a family, he lives in a certain society, in a certain state. What he has to do depends on what his place is, what his function is, and that all comes from his station in the organism (173).

Although Bradley expresses his views regarding the nature of the self in a way which

is elegant, forceful, and endlessly quotable, there are nevertheless some serious problems with his account.⁸¹ First of all is his continued insistence that he is not appealing to metaphysics to support his claims about morality, and is instead appealing to the facts and the common moral consciousness.⁸² The truth is that in *Ethical Studies* Bradley is not offering any *arguments* for his metaphysical principles. He nevertheless does *assume* a rather exotic, or at least from our perspective, an extremely controversial set of metaphysical propositions. To the extent that Bradley relies on the notion of an infinite whole of which we as individuals are mere heartbeats in the system, his conclusions are on rather shaky ground. This is not to say that it is inconceivable that the idealist metaphysics might turn out to be true; however, given the *prima facie* implausibility of Bradley's metaphysics, the burden of proof seems to be on the adherent to idealism to make the case for its metaphysical assumptions. Although Bradley did later work out his metaphysical system in considerable detail in *Appearance and Reality*, that system does not seem to have taken hold within the general philosophical community. Until such an idealist metaphysics comes to be more widely accepted than it presently is, those of us who are skeptical of it can safely assume that any appeal to it in support of ethical conclusions is highly suspect.⁸³

⁸¹ Broad complained about the relative influence of Bradley and Sidgwick, suggesting that the greater influence of Bradley (at least up to the early 20th Century) was due to his superior literary style and was in spite of Sidgwick's superior philosophical ability (Broad: 144).

⁸² See, for example, 65, 165-66, 247. See also Nicholson 52n26.

⁸³ The story of how the sort of idealism championed by Bradley fell out of favor in the 20th Century is no doubt complex. But Moore and Russell are in any event central figures in its decline. Moore's famous paper, "The Refutation of Idealism" was published in 1903 and influenced many, including Russell (Hylton: 117, 245). Russell's doctrine of logical atomism and his account of the foundations of mathematics can also be thought of as systematic attempts to work out an alternative to late 19th Century British idealism. See Hylton for an extended discussion of Russell's rejection of idealism. Also, logical positivism, with its

Moreover, on an intuitive level, Bradley's claims about the essential social nature of the self seem false. It is important to note that Bradley is not merely claiming that humans are essentially social beings, or that certain general social dispositions—e.g., a disposition toward friendship or toward certain kinds of social cooperation—are essential to humanity. He is instead making the more radical claim that one's essence is a matter of one's particular community; my essence is constituted by the particular set of attitudes, beliefs, values, and relations one typically associates with a late 20th, early 21st Century American. However, this socialized conception of the self seems to conflict with our considered judgments about the sorts of alterations a person can undergo without passing out of existence.

It would seem that I can give up many of the relations, attitudes, beliefs, etc. that are supposed by Bradley to be part of my essence without passing out of existence. Bradley himself admits that an Englishman can cut his ties with England without ceasing to exist, and that—even if the rest of the nation perished, he would nevertheless survive (166). However, if this is true—and it does seem to be true—it is hard to see how it is that these relations tying the Englishman to English society are essential to his being. At most, we would have to admit that the content of one's personality is largely a function of one's education and environment, and that there is a *sense* in which one's identity is a function of this content. Thus, we might be willing to say that if I had been raised by nomads in Afghanistan I would have turned out to be, in *some* sense, a different person. Similarly, if a neuroscientist altered my brain chemistry so radically that I suddenly took on all of the traits and dispositions of an Afghan

dismissal of much of traditional metaphysics as cognitively meaningless, no doubt played a large role in helping to consign idealism to the dustbin of history. However, there are some holdouts. For example, Sprigge offers an example of a fairly recent defense of absolute idealism.

nomad, then we might very well say that I had ceased to exist. However, we would still have to ask what precisely we mean by such a claim, since to many it would seem perfectly sensible to say that *I* could have been raised by Afghan nomads, or that *I* could be transformed into an Afghan⁸⁴. Thus, it is not at all clear that, when we say that the neuroscientist's guinea pig is a different person, we are really committed to the proposition that his personality is essential to his being. Our saying that he is a different person might just be a picturesque way of saying that he has changed so much that he is hardly recognizable.

The best we can do for Bradley's position, I think, is to assume that personal identity is a function of the beliefs, values, attitudes, and relations which seem to make up one's personality. But now the problem for Bradley is to specify precisely how much of the personality has to change before one has not only altered, but ceased to be. Given that there is an almost infinite number of small alterations the expatriate Englishman can make to his personality without passing out of existence, one suspects that none of these traits is essential to his being. In other words, on this understanding of the self, it is extremely difficult, if not impossible, to point to any particular trait or relation and say that it is particularly central to one's being. The most that we can say is that, if one is interested in continuing to exist, one should avoid sudden, wholesale changes to one's personality, because if one changes too

⁸⁴ This example is complicated by the fact that we might think that bound up with the concept of being an Afghan is something like actually having lived in that country, so that the neuroscientist cannot *literally* turn someone into an Afghan; for present purposes, we can simply stipulate that what we mean by an Afghan is roughly someone with the various mental states and behavioral dispositions typically associated with someone with the right kind of roots in the relevant environment. It might help to substitute a category that is more ideological in nature, such as Islamic Fundamentalist or 12th Century Christian. Cf. Locke's famous discussion of the prince and the cobbler in his *Essay* (Book II, chapter 27).

drastically, too suddenly, then one ceases to exist. While this may or may not be a plausible account of personal identity, what is important for our purposes is that it seems incapable of underwriting Bradley's claims about my station and its duties. It seems undeniable that, on this view, one can easily justify failing to fulfill one's function within the social organism, since even a systematic and repeated shirking of one's duties seems compatible with maintaining the basic integrity of one's personality.

Even if we put these objections aside and accept Bradley's claims about the nature of the self, along with his claim that the fulfillment of the duties associated with one's station in society is constitutive of self-realization, we still have an account of the good and self-realization which is implausible and unattractive. For it is difficult to see how the fulfillment of my station and its duties could be anything more than merely *partly* constitutive of my self-realization. Surely, there is more to self-realization or a good life than the fulfillment of such duties. Indeed, it is commonplace that our duties often require some form of self-sacrifice. The fulfillment of my duties may very well be painful, or may require that I give up other goods, such as increased knowledge or the opportunity to pursue an artistic or athletic endeavor. At the very least, it seems clear that My Station's account of self-realization must be supplemented by other goods, or at least by a richer account of self-realization. Moreover, it will not be at all surprising if these other goods or aspects of self-realization turn out, in many cases, to conflict with the demands of my station in the social organism.

A favorite analytical tool of moral philosophers involves the segregation of various issues in ethics—e.g., metaethics is segregated from normative ethics, questions about the good are separated from questions about the right, and so on. Although this can often be a helpful way of dealing with such issues, it is a methodology seriously at odds with the

approach of philosophers such as Bradley. Bradley seems to willfully blur the lines between these various issues, and it is very difficult to come to terms with his view unless one is prepared to consider a question in his terms. For example, I am mainly concerned with Bradley's account of self-realization as an account of the *good*. However, he offers self-realization as an account, not just of the good, but of morality in general. A large part of the point of Bradley's approach is the blurring of the boundaries between the good and the right, or between self-interest and morality. Thus, in order to evaluate Bradley's account of self-realization, we should be prepared to consider self-realization as an account of the *right* or other-regarding morality. And it seems clear that there is often much more to fulfilling one's moral obligations than simply living in accord with the duties associated with one's place in the social organism.

Bradley himself is aware that *My Station* offers only a partial account of both the good and the right. For Bradley, the realization of the good self has three distinct components. The most important is constituted by *My Station*:

[The reader] will find, if we do not mistake, that the greater part of [self-realization] consists in his loyalty, and according to the spirit, performing his duties and filling his place as the member of a family, society, and the state. He will find that, when he has satisfied the demands of these spheres upon him, he will in the main have covered the claims of what he calls his good self. The basis and foundation of the ideal self is the self which is true to my station and its duties (220).

However, there must be more to morality than fulfilling such duties. Imagine, for example, a member of the *Waffen S.S.* who thinks that he *must* brutalize Jews in Poland because it is his job; he conceives of the persecution and murder of innocent civilians as a large part of what

constitutes his function within the social organism.⁸⁵ Virtually all of us would say that such a man has a monstrously narrow conception of his duties, and with this Bradley would seem to agree; according to Bradley, it is necessary to remark that the community in which [one] is a member may be in a confused or rotten condition, so that in it right and might do not always go together (203), and considerations of my station and its duties have to be supplemented by other considerations. As Bradley puts it, you cannot confine a man to his station and its duties (204).

Bradley deals with this problem of the inadequacy of My Station vis-a-vis our judgments about the good and the right by introducing what he calls Ideal Morality. Ideal Morality has two aspects: a social ideal and a non-social ideal. The non-social ideal, which involves the realization of what Bradley calls the non-social ideal self, involves the single-minded pursuit of knowledge and beauty, and is a possible response to the complaint that there must be more to a good life than merely the fulfillment of one's function in the social organism.⁸⁶ In response to the concern that there must be more to morality than the mere

⁸⁵ Although the causes of the Holocaust are undoubtedly complex, at least part of the story has to do with a strong sense of social duty among German soldiers, administrators, and police. About the Nazis' decision to begin the policy of extermination of Jews in Poland, two historians have the following to say: The systematic dehumanization of the Jews by the Nazis since 1933, culminating in the Polish ghettos, helped to remove the moral inhibitions which would otherwise have prevented the contemplation let alone the implementation of a policy of extermination. In addition there was a widespread mentality of professionalism in which the efficient performance of a professional function became the overriding consideration and moral criteria were reduced to secondary virtues such as duty and loyalty (Noakes and Pridham: 1102-1103).

⁸⁶ It is unclear why Bradley doesn't simply argue that, for some people (viz., artists and scientists), their station in society requires a deep commitment to the acquisition of knowledge or to the creation of beauty. Presumably, it is the single-minded nature of the work which makes it simply implausible to think that one can live the life of the great artist or scientist while still finding time and energy for the fulfillment of the duties of family and country. At any rate, Bradley makes it clear that he thinks there is a kind of self-realization to be found in scientific inquiry and artistic endeavor which does not reduce to, and can often

fulfillment of the duties associated with my station in society, Bradley introduces the notion of an ideal social self. It is an ideal because, unlike the duties associated with My Station, it is not rooted in any actual, existing social organism. We see bits and pieces of the social ideal exemplified in the actions, beliefs, and commitments of certain people, some of them members of our own community who are blessed with a heightened moral sensitivity or imagination, some of them from outside our community (221-22). As Bradley puts it: ...we seem first to see in some person or persons the type of what is excellent; then by the teaching, tradition, and imagination of our own and other countries and times, we receive a content which we *find* existing realized in present or past individuals, and finally detach from all as that which is realized wholly in none, but is an ideal type of human perfection (221).

Bradley is frustratingly vague about the specific content of the social ideal, often appearing satisfied merely to assert that there is such an ideal which we are vaguely aware of, leaving it to the reader's imagination to determine just what it consists in. What Bradley *is* clear about is that the social ideal goes beyond my station and its duties, and that while it is an ideal to be striven for, it cannot be fully realized or attained (220). One possible interpretation of the social ideal is to take it as a demand for what moral philosophers today would call full agent-neutrality. This would fit with Bradley's views regarding the relation between individuals and society as a whole. In his influential critique of utilitarianism, John Rawls focuses on the agent-neutrality of that view, complaining that the principle of rational choice appropriate for one person is extended to society as a whole, as if the relation between society and its individual members could be modeled on the relation between time-slices of an individual's life and her life as a whole. According to Rawls, utilitarianism's view of social

conflict with, the self-realization to be found in my station and its duties (214, 222-23).

cooperation is the consequence of extending to society the principle of choice for one man, and then, to make this extension work, conflating all persons into one through the imaginative acts of the impartial sympathetic spectator (Rawls 1971:27). However, given Bradley's metaphysical assumptions, according to which the relation between the individual and society is basically a part-whole relation, he does not even need the impartial spectator's acts of sympathetic imagination. Bradley's views regarding the metaphysics of the individual and society give him his own reasons not to take seriously the separateness of persons.

Moreover, there is some textual evidence that Bradley intends the social ideal to be understood ultimately in agent-neutral terms. In discussing the ideal self, Bradley says the following:

So far we have seen that the self is identified with pursuits and activities as ends to be gained by it, but further it is interested in persons and causes which stand in no direct relation to its personal activity. Apart from anything which it does or has to do, it feels its will affirmed or denied in the success or failure of that which its own action has not to bring about. This result is a mere continuation of the process which drops everything subjective, everything which concerns only me in particular, out of the content of the end, and subordinates my aims to general heads, until on the one hand the mere objective content of the ends, apart from the ideal of my activity, is felt as the affirmation of my will, and on the other hand those ends are brought into a harmony, over which presides what, for shortness sake, we may call the ideal (292).

Although it is difficult to know how this apparently agent-neutral social ideal is supposed to relate to other aspects of self-realization such as my station and its duties and the pursuit of truth and beauty e.g., are we supposed to be moving toward complete agent-neutrality, or can we be content to look beyond these more provincial demands and take on the more objective point of view merely from time to time or only in certain situations? it does

nevertheless seem to be the case that part of our self-realization can be found in the promotion of agent-neutral value.

One way of reading Bradley is to take him simply as being ambivalent about the relation between the demands of My Station and those of the social ideal which I have been interpreting as agent-neutral. Sometimes he wants to say that the agent-neutral social ideal is just one relatively small part of our self-realization, while other times he wants to say that full agent-neutrality is what we should all be striving for. For example, in the essay on *Ideal Morality*, Bradley tells us that it is a moral duty to realize everywhere the best self, which for us in this sphere is an ideal self; and asking what morality is, we so far must answer, it is coextensive with self-realization in the sense of the realization of the ideal self in and by us (219). He then immediately identifies the ideal with (1) My Station and its Duties; (2) the social ideal which goes beyond My Station; and (3) the pursuit of truth and beauty (219-24). However, later in *Selfishness and Self-Sacrifice*, Bradley seems to suggest that the realization of the moral ideal requires us to leave the familiar duties of (1) and (3) behind, and find ourselves in the good of the whole:

The man's self is now wrapped up in the general progress of good, his will is so far by habituation become one with the ideal; and in the realization of that, whether by himself or others, he finds a permanent and everlasting source of pleasure; a cause which brings indeed its own pains with it and, in the absence of faith, can do much to sadden, but in which alone he finds his true self affirmed, and affirmed apart from his private success or failure (292).

Here, the suggestion seems to be that we find our true selves in the good of whole, and achieve our own good or self-realization when we subordinate our own private will to the will of the organism of which we are merely a part. Bradley is perhaps more explicit about this a few pages later:

[The good self] knows itself at first as the will which, against the temptation of the bad, wills in its acts, and wills its acts as, the will of a superior outside itself, whether that be a person or a tribe. The higher will is here felt, but not yet known, to be also the will of the obeying self; and the process of development, whether in morals or religion, has for its result the end where this higher will is known as the true will of the self, where law ceases to be external and becomes autonomy, and where goodness, or the identity of the particular will with the universal, is only another name for conscious self-realization (300-01).

Here, as in the previous passage, the true self is identified with the universal will (i.e., the will of the social organism), and it is suggested that one finds ultimate self-realization by submerging oneself in the social organism.⁸⁷

There seems to be a good deal of tension in Bradley's account of self-realization. At times he tells us that self-realization is multifaceted, including my station and its duties, a further social ideal, and the pursuit of knowledge and beauty. At other times, Bradley suggests that true self-realization involves an identification of oneself with the universal will and a losing of one's private self in the consciousness of the social organism. Perhaps we need to appeal to Bradley's Hegelian, Dialectical method to account for this tension. Bradley makes it clear that he is some manner of moral relativist (189-92). He sees things in terms of an inevitable march of history, and thinks that the social and moral ideal of each place and epoch is appropriate for it (192). It is quite possible that Bradley thinks that our current ideal consists in My Station, the pursuit of truth and beauty, and something which goes beyond the duties of my station in society, the content of which it is impossible to give a precise account of. The more demanding ideal, the one which seems to call for full agent-neutrality and the identifying of the individual will with that of the universal will, is perhaps something which

⁸⁷ See also p. 244, where Bradley refers to the bad self as a false self.

will be appropriate to a future society, and is perhaps an ideal to which we are inevitably drawn.

While this interpretation of Bradley may resolve some of the tension in his thought about self-realization, it has significant costs. First, it puts us back in the position of wondering what could be the content of the social ideal discussed in *Ideal Morality*. Since we have given up on the idea that the supplemental duties of ideal morality consist in a demand for agent-neutrality, the notion has now been stripped of the content which our original interpretation had given it. More importantly, we are left wondering how the various aspects of self-realization are supposed to relate to each other. It does not help to tell us that some future state of man will be characterized by full agent-neutrality and the identification of individual wills with the universal will. For *ex hypothesi* we do not occupy such a state; we instead occupy the state where self-realization is to be found in three distinct arenas: the duties of my station; a further, but unspecified, social ideal; and the practice of science and art. Unfortunately, Bradley gives us no indication of how we are to weigh these various aspects of self-realization against one another should they conflict. Moreover, Bradley has given us no reason to think that they will not conflict; indeed, it seems commonplace for the duties associated with my station in society to conflict with these other goods. What we need for the resolution of possible conflicts is a conception of self-realization which allows us to adjudicate cases where fulfilling the duties associated with my station in society will not best promote beauty and knowledge or be in accord with the (unspecified) social ideal, or where the duties of the (unspecified) social ideal conflict with my station in society or with the pursuit of art and science. Bradley offers no such conception of self-realization.

This problem of conflicts among various aspects of self-realization is made worse by

some of the things that Bradley says about pleasure. As we saw in the previous chapter on Hurka's perfectionism, it can be a challenge for a perfectionist to account for our considered judgment that certain states, such as pleasure, are intrinsically good. The status of states such as pleasure and pain is also a problem for Bradley. At times, Bradley seems to want to solve the problem by emphasizing that we often feel pleasure when we achieve self-realization, as he understands it. Although it is true that we do often feel a certain kind of pleasure at this sort of self-realization, this is not necessarily the case. Indeed, the fulfillment of the duties associated with my station in society could very well be quite painful, or cause me great distress or anxiety. The same seems to be true of the single-minded pursuit of knowledge and beauty; there's no reason to think that the life of the dedicated scientist or artist will necessarily be a particularly pleasant or even happy or contented life. Moreover, although there is, as we have seen, some question about the ultimate content of the social-ideal aspect of self-realization, it seems likely, whatever its content, that its fulfillment will often require a sacrifice in pleasure, contentment, or happiness.

Bradley also says things at various points which seem to commit him to admitting that there are goods which are distinct from self-realization, as he understands it. At the end of *Pleasure for Pleasure's Sake*, where Bradley makes the case against hedonism, he admits that happiness (though, not pleasure) is indeed an end in itself (125). Now Bradley might be able to get away with saying that happiness is the final end without admitting a good distinct from self-realization, since there is a long tradition associated with Aristotle and his followers of understanding happiness in non-subjective terms.⁸⁸ However, one must at least be wary of the fact that the rhetorical force of such claims often trades on the implicit

⁸⁸ See for example Kraut's *Two Conceptions of Happiness*.

assumption that happiness must have at least *some* subjective dimension. Later in *Ethical Studies*, Bradley admits that happiness and virtue do not necessarily go together:

If what is meant be this, that what is ordinarily called virtue does always lead to and go with what is ordinarily called happiness, then so far is this from being verifiable in everyday experience, that its opposite is so; it is not a fact, either that to be virtuous is always to be happy, or that happiness must come from virtue (318).

Perhaps the term happiness is ambiguous, sometimes having an objective sense and at others a subjective one. Maybe when Bradley tells us that happiness is an end in itself he is employing the objective usage of the term, and later, when he denies that happiness and virtue have to go together, he is adopting the subjective sense of happiness. The problem with this line of thought is that it ignores the fact that when philosophers say things like, We all desire happiness or Happiness is an end in itself or Happiness is *the* final end, the plausibility of such claims very often presupposes the more subjective sense of the term. While things may have different for Aristotle, so that he could identify happiness or eudaimonia with rational activity in accord with virtue, for us, Bradley's claims about happiness being an end in itself must have at least a strong subjective component if they are to be as obviously plausible as he seems to assume.

Around the point where Bradley is telling us that happiness is an end in itself, he also makes the familiar anti-hedonist move of saying that while pleasure might be *a* good, it is not *the* good (125). While this claim is familiar and, and to most people (I would think) plausible, it is not at all clear how it fits in with a perfectionist account of the good. While such a claim might offer a defense against a view which attempts to draw a line from the goodness of pleasure to hedonism, it does not seem to be of great assistance in a positive defense of *perfectionism*, since unless we can give a perfectionist account of the goodness of pleasure, it

would seem to lead us to something like value pluralism. Since Bradley seems unable to explain the value of pleasure in terms of self-realization as he understands it, it would seem that his admission that pleasure is a distinct good leads to a version of pluralism about the good, according to which the good consists in (at least) self-realization and pleasure.

The notion that self-realization and pleasure can be reconciled is further undermined by some claims about selfishness and self-sacrifice which Bradley makes in his last essay. Although he makes a half-hearted suggestion that self-sacrifice and selfishness are ultimately the same (251), Bradley nevertheless feels compelled to give a more plausible account of selfishness. In the most general terms, what constitutes selfishness is thinking only of yourself (274). According to Bradley,

The selfish man, so far as he is selfish, has objects of desire which are not subordinated to any principle higher than his private satisfaction. If you ask what is the general end which includes his ends, you can point to none; but you find that he treats all objects as means, that he cares for none in itself, but will sacrifice any with readiness; and when you inquire what is common to them all, you find that they minister to his personal comfort; this comfort being a certain quantum of the pleasant and the absence of pain, which satisfies him, and which he either consciously aims at or unconsciously uses as a measure of all objects of desire (274-75).

Indeed, in his discussion of what selfishness is, Bradley mainly emphasizes states such as contentment, satisfaction, pleasure, and the absence of pain (274-76). Although this leaves open the question of the relative importance of states such as pleasure and contentment in Bradley's view, it nevertheless seems undeniable that they are being admitted as genuine goods. Otherwise, it is hard to see why we should think of their pursuit as *selfish*. The selfish man is someone who pursues his own private good at the expense of the good of others, usually in a case where this involves the agent's overestimation of the importance of his own

good compared to the good of others.⁸⁹ Moreover, unless we think that the agent's pleasure is a genuine good for him, it is hard to understand why we typically think that others are justified in being *resentful* of his pursuit of his own pleasure at the expense of others.

In this section I've been discussing Bradley's self-realization version of perfectionism. One of the features of self-realization accounts of the good or morality is that they tend to blur the line between an agent's good or well-being and the moral goodness of the agent's actions or character. I've suggested that this is an attractive feature of such views, since it offers an account of the authority of moral considerations or reasons. Unfortunately, Bradley's version of self-realization is ultimately untenable. In the end, Bradley presents us with a fairly disparate list of goods to which he attaches the label, self-realization. We are told that our selves are realized in the fulfillment of those duties associated with our function in the social organism. We are then told that there are further duties, the specific content of which is left somewhat vague, which go beyond the duties of one's station in society, but which nevertheless are also constitutive of our self-realization; I've interpreted these further duties as involving taking on an objective, agent-neutral point of view. We are also told that one can find self-realization in the pursuit of knowledge and beauty, à la the dedicated scientist or great artist. Finally, pleasure and related states seem to demand a place on the list of goods, and Bradley says some things which suggest that he agrees.

Much of the story that Bradley tells about self-realization is plausible as an account of our other-regarding moral considerations. It seems true, for example, that there are certain obligations which come with the occupying of a certain role in society. Also, it seems to be

⁸⁹ Cf. Bradley p. 276, footnote. In this note, which was added for the second edition of *Ethical Studies*, Bradley says that self-realization can itself be selfish if it means undue subordination of others' welfare to one's own.

the case that these duties associated with one's station in society are not exhaustive of our obligations; sometimes, we are required to go beyond our station and its duties, perhaps even to the point of taking on an objective, god's eye view of things. However, Bradley's attempt to account for these obligations in terms of self-realization is unsuccessful. Without a commitment to the metaphysics of idealism and an implausible, socialized conception of the self, one is left with no reason to think that one's own good or well-being is constituted by the fulfillment of such duties. Similarly, while many would share Bradley's conviction that knowledge and beauty are intrinsically valuable, and that the life of the scientist or artist is a noble calling, it is hard to see why these have a special status in Bradley's scheme.

It is difficult to come away from Bradley's *Ethical Studies* without the conviction that much of Bradley's thinking about morality and well-being is confused, muddled, or ultimately incoherent. Some of this is, no doubt, a false impression created by Bradley's Dialectical method. However, it is probably also the case that the Hegelian methodology is often a mask for genuine inconsistencies in Bradley's thought. At times, Bradley suggests that there is no tension between self-sacrifice and selfishness, and that we find self-realization through doing for others; at other times, he admits that selfishness involves the pursuit of one's own good at the expense of others. At one point, Bradley tells us that we are all, in some sense, motivated by self-interest; at other times he emphasizes the ways in which we can come to desire the good of others as an end in itself. At times, Bradley suggests that morality is really not that demanding, and that we can fulfill our obligations by focusing on the duties of our station; at other times he suggests that morality requires that we take on an objective point of view, and treat our own desires or welfare as being of no more importance than that of others. In the end, it is difficult to avoid the conclusion that Bradley is trying to unify that which cannot be

unified, or reconcile that which is irreconcilable. Bradley has given us no reason to think that morality and self-interest are not often seriously at odds with one another, and he has failed in his attempt to give a unified account of morality, self-interest, and the good in terms of self-realization.

III. Green on Self-Realization

In the previous section, we saw that one of the main problems with Bradley's account of the good in terms of self-realization is its reliance on a set of metaphysical assumptions which, if not untenable, are at least extremely controversial. Since the metaphysical and epistemological views of Bradley and Green are usually taken to be very similar,⁹⁰ this would seem to make Green's account of self-realization *prima facie* unpromising as well. Although Green does devote a significant amount of space in the *Prolegomena* to an explication and defense of the metaphysics and epistemology of idealism,⁹¹ it is possible to read Green as relying on a less controversial set of metaphysical views. David Brink has made considerable efforts toward a reconstruction of Green's view which does not appeal to skepticism about the metaphysical status of individuals or to the related conception of persons as aspects of an interpersonal organic unity (Brink 1997b: 133). Brink attributes to Green a conception of the self which is very similar to that which is found in the eudaimonistic tradition (especially as exemplified by Plato and Aristotle), echos of which can be found in the views of modern philosophers such as Locke, Butler, Reid, and Kant. Instead of addressing the issue of whether Brink's reconstruction is an accurate interpretation of Green's such exegetical issues

⁹⁰ For a helpful discussion of the metaphysical and epistemological views of Bradley and Green see chapters 1 and 2 of Hylton.

⁹¹ See especially Book I.

are too complex to be addressed in such a limited space, and are of limited relevance to my purposes here anyway. I shall simply assume that Brink's reading of Green is basically right and concentrate on its plausibility and attractiveness as an account of the good.⁹² I shall argue that it is so implausible and unattractive that it ought to be rejected in favor of pluralism about the good.

In discussing Bradley, we found that it was difficult to separate his thinking about self-realization as an account of morality from his thinking about it as an account of the agent's good or well-being. This was not that surprising since a large part of his project seems to involve the blurring of the boundary between morality and self-interest. Brink's reconstruction of Green, by contrast, makes it easier for us to keep the two issues separate. Green's account⁹³ of the nature of the self allows us, for the time being, to set aside issues about the content of the agent's good and to focus instead on how it is that the good of others can be constitutive of the good of the agent. Green's conclusions about the relation between the good of the agent and the good of others—at least on Brink's reading of Green—depends on views about the nature of the self and the identity of persons, and seems to be largely independent of the question of the content of the good. Since one of the most attractive features of Green's view is that it offers an alternative to Bradley's notion of an interpersonal organic unity, I will begin by discussing Green's account of how it is that we come to have our interests extended and why he thinks that the good of others is a constituent of our own

⁹² It is in Brink 2003 that he is most explicit about tying his account of the good to that of Green. However, the roots of the Green-inspired view can be found in Brink 1989, especially chapter 8, §2. Also relevant to Brink's take on Green's perfectionism, in particular his account of corporate agency and the common good, are Brink 1997a and 1997b.

⁹³ I will often refer to Brink's reconstruction of Green simply as "Green's account" for the sake of convenience. From here on, whenever I refer to Green or Green's account the reader should assume that I mean Brink's interpretation of Green, unless I indicate otherwise.

good. I will then move on to a consideration of Green's account of the content of the good.

We should begin by asking a fundamental question: What is the self? Although Green does not—as far as I can tell—explicitly refer to John Locke's famous account of the self and personal identity, Brink nevertheless thinks that Locke's view is the most natural starting point for coming to a plausible account of the self.⁹⁴ On Locke's view, the concept of a self is intimately connected to that of a person. Indeed, for Locke 'person' is just another name for the self:

Person, as I take it, is the name for this *self*. Where-ever a Man finds, what he calls *himself*, there I think another may say that is the same *Person*. It is a Forensic Term appropriating Actions and their Merit; and so belongs only to intelligent Agents capable of a Law, and Happiness and Misery (*Essay* Book II, chapter 27, section 26).

Locke also says that 'person' stands for 'a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places' (II, 27, 9). In short, a person is a self-conscious entity which persists over time, is capable of exercising deliberative control, and is thereby responsible for its actions. Brink places Locke's conception of the self within a tradition which identifies the self as that aspect of us which exercises deliberative self-control. Two of the most prominent members of this tradition are Plato and Aristotle, both of whom equate a person with the controlling or rational part of the soul—that is, with the understanding (Brink 1997b: 126-33).

How exactly does Green figure into this tradition which identifies the self with the understanding? For Green, the self is characterized by two types of responsibility: epistemic and practical. To say that an agent is epistemically responsible is to say that she has the

⁹⁴ See Brink 1997b: 136-138, 2003: 67-69.

capacity to distance herself from her doxastic impulses; that is, the *epistemically* responsible agent is able to critically evaluate the various appearances she is presented with, whereas a child or an animal merely accepts appearances in an uncritical manner. Similarly, the *practically* responsible agent is able to distinguish between the intensity and authority of his desires (Brink 1997b:134). An animal or a child, by contrast, seems merely to be a slave to its desires; it is incapable of distancing itself from its present desires or of forming a conception of its overall good, according to which it might regulate its behavior. I, on the other hand, as a rational agent, am able to critically evaluate my present desires in light of a conception of my own good. Thus, while the desire to eat Oreo cookies may be extremely strong or intense, I can form a judgment according to which my good consists in the frustration of the present desire to eat cookies – perhaps because I conceive of my good as consisting partly in physical health or in a svelte appearance. A person is someone who regulates her actions and beliefs through rational deliberation about the appropriateness of her present desires and beliefs, and that part of her that exercises such deliberative and doxastic control is the self.

Although Brink's interpretation of Green does not rely on the notion of an interpersonal organic unity to explain how the good of others can be a constituent of the agent's good, he does nevertheless have a metaphysical story to tell about how there can be such a thing as corporate agency. However, Brink's story about corporate agency appeals to views about the nature of persons and the persistence of the self which, in contrast to the metaphysical and epistemological assumptions of idealism, are well within the mainstream of current thinking about personal identity. Brink understands the self in terms of psychological continuity and connectedness, and he is certainly not alone in conceiving of the self in such

terms: David Lewis, Derek Parfit, Anthony Quinton, and Sydney Shoemaker are all examples of contemporary philosophers who have accepted such a position.⁹⁵ Brink appeals to psychological continuity and connectedness to explain how the various temporal stages of a typical life are unified, and then argues that different individuals can be related to one another in an analogous way. Thus, although Bob and Mary might be separate individual human beings, if the right psychological relations hold between them, the boundary between Bob-the-person and Mary-the-person can get blurry, so that a kind of corporate agency can hold between them.⁹⁶

An example illustrating the analogy between identity within a life and identity among lives would perhaps be helpful. Take Bob for instance. There are various time slices or temporal instances of Bob. At time *t*1 Bob forms the intention to travel from his home in San Diego to Los Angeles for a Dodgers game six months hence. Bob does indeed go to the game six months later, but how do we know that the Bob who goes to the game is identical to the Bob who formed the intention to go six months earlier? Our intuition is that the two Bobs are one in the same, and what underlies this intuition is our notion that there is a self which persists through the various stages of Bob, stretching from the forming of the intention to go

⁹⁵ Specifically, these philosophers all seem to accept what Parfit calls psychological reductionism, according to which claims about personal identity are analyzed in terms of claims about psychological connectedness and continuity. It should be noted, however, that one can accept a psychological account without accepting reductionism; one might instead think that claims linking personal identity with certain psychological states or phenomena are akin to synthetic identity claims, such as heat = mean molecular kinetic energy. Mark Johnston describes psychological reductionism as the current dominant account of personal identity and cites Parfit, Lewis, Quinton, and Shoemaker as prominent adherents of the view (Johnston: 61).

⁹⁶ Brink's views on personal identity and agency can be found in a number of places. Both *Self Love and Altruism* (especially 136-43) and *Perfectionism and the Common Good* (63-69) contain brief discussions. For a somewhat lengthier discussion, see Brink's *Rational Egoism and the Separateness of Persons*.

at time $t1$, to the invitation to Come-along Carl at time $t2$, to the buying of the tickets at $t3$, to the borrowing of the sister's car at $t4$, to the attending of the game at $t5$. But how do we cash out our intuitive notion that what makes Bob-at- $t1$ identical to Bob-at- $t5$ (and all the other Bob's) is a persisting self? One answer is that the beliefs, intentions, and actions of the various Bob's are all connected in the right sort of way so that they can all be thought of as instances of the persisting self. For example, the fact that the action of going to the game at $t5$ is causally dependent on the forming of the intention at $t1$, along with all the other actions and intentions along the way, makes it the case that the various temporal stages of Bob are just instances of the same person. As Brink puts it,

what makes persons at different times the same person and, hence, what unites different parts of a single life is psychological continuity. A series of persons is psychologically *continuous* insofar as contiguous members in the series are psychologically well connected. A pair of persons are psychologically *connected* insofar as the intentional states (e.g., beliefs, desires, and intentions) and actions of one are causally dependent upon those of the other. Of particular importance, given our views about persons, are *deliberative* connections that hold among actions, intentions, and prior deliberations in the deliberate maintenance and modification of intentional states and in the performance of actions that reflect these prior deliberations (1997b: 138).

Instead of holding that what unifies the various stages of Bob is something like a Cartesian ego or otherwise mysterious subject of experience, Brink argues that it is the interrelations among the various intentions, beliefs, desires, and actions of the various stages of Bob which enable the relation of identity to hold among the various temporal instances of Bob. This is the sense in which Bob-the-person persists over time.

Given this account of identity within a life, it is not hard to see how an analogy can be drawn with the relation between different individual human beings. Imagine that Bob and

Mary have a very close relationship for instance, they have been married for twenty years. It would not be at all surprising if the various intentional states and actions of Bob were related to the intentional states and actions of Mary in much the same way that the intentional states and actions of Bob-at-time-*t1* are related to those of Bob-at-time-*t5*. Thus, if there is sufficient psychological connectedness and continuity between Bob and Mary, then the boundary between Bob-the-person and Mary-the-person can become quite blurry. Although the degree of continuity required for Bob and Mary to be one person may be extremely rare or even non-existent, Brink argues that this *type* of psychological continuity is commonplace:

Though I am normally most strongly continuous with myself in the future, I can be psychologically continuous with others with whom I interact psychologically. Interpersonal, as well as intrapersonal, continuity is quite common. Interpersonal connections and continuity can be found among intimates who interact on a regular basis and help shape each other's mental life; in such relationships, the experiences, beliefs, desires, ideals, and actions of each depend in significant part upon those of the others. We can see this in the familial friendships that Plato, Aristotle, and Green all take as their model. Parents make plans for their children that affect their children's actions, opportunities, and experiences; they impart information and teach skills; they make suggestions, act as sounding boards, and set limits. In these and countless other ways, parents help shape their children's faculties, experiences, beliefs, desires, values, opportunities, and goals. Similar relations hold among spouses and friends who share experiences, conversations, and plans. They can also be found, to a lesser extent, among partners in cooperative ventures where the deliberations, desires, plans, and expectations of each are formed together and conditioned by each other (1997b: 141).

It is important to emphasize that, on this view, continuity admits of degree, and it may seldom, or even never, be the case that there is sufficient psychological continuity between two different human beings such that they constitute one person in the same way that two contiguous time slices of Bob constitute one person. Nevertheless, if we assume that concern

for the welfare of our future selves depends on the identity relation as understood in terms of psychological continuity, and if we assume that the *degree of concern* should be a function of the *degree of continuity*, then it will be natural to conclude that we should be concerned with the good of others we are sufficiently psychologically continuous with in the same way as, and for the same reason that, we are concerned with the good of our more distant future selves. For example, if Bob-at-time- $t1$ is more continuous with Mary-at-time- $t1$ than he is with Bob-at-time- $t101$, then at $t1$ Bob should have greater concern for Mary than for himself at $t101$. Moreover, if we keep in mind that continuity, and hence identity, can admit of degree, then it will seem considerably less outré to claim that there can be a kind of corporate agency holding between or among distinct human beings. Thus, we see how Green can get at least some degree of corporate agency without relying on extremely obscure and controversial notions of an interpersonal organic unity.

Although this account of limited corporate agency is controversial, it is nevertheless a plausible enough account of how it is that the good of others can be a constituent of the good of the agent that we can assume its soundness for the sake of argument and move on to a discussion of the content of Green's conception of the good. For our present purposes, it is sufficient to show that there is a viable alternative to the traditional idealist notion of an interpersonal organic unity, and to note that this alternative account of corporate agency would meet with the approval of many prominent contemporary philosophers who have concerned themselves with questions about the nature of persons and the persistence of the self; I will eventually make some criticisms of this view, but I want to put that off for a bit. A second thing to note about this account of corporate agency is that it leaves open the question of what the good of an agent consists in. If hedonism is true, and pleasure is the only thing

worth pursuing for itself, then I should work to promote the pleasure of those who are sufficiently psychologically continuous with me; if a version of the desire-satisfaction view is correct, then I should work to see that the relevant desires are satisfied. If some form of perfectionism is correct, then I should promote their perfection or excellence. Thus, we now need to turn to Green's account of the content of the good.

What precisely is Green's account of the good? Clearly, he is some sort of perfectionist, and he equates the good or perfection with self-realization. But what exactly is self-realization? One way of understanding Green's view is to focus on his conception of the self. Recall that, for Green, the self is that part of us which makes us epistemically and practically responsible agents. We are epistemically responsible because, unlike young children and brutes, we do not simply assent to our appearances. We are not slaves to our doxastic impulses; we evaluate evidence for our beliefs, and (at least to the extent that we are epistemically rational) modify our beliefs accordingly. Similarly, we are practically responsible because we are not slaves to our desires. We are capable of distinguishing between the intensity and the authority of our desires, so that even though the desire to (say) take revenge on an enemy might be quite intense, unlike brutes and small children, we do not have to act on that intense desire. We can instead recognize that our desire to avoid a potentially disastrous confrontation, though less intense, has greater authority, and to the extent that we are practically rational, we are able to regulate our behavior according to our judgement of what is best all things considered.

For Green, perfection consists in the development of the self. The self is that part of us that exercises deliberative control (both epistemic and practical). A natural way of reading Green's perfectionism conceives of self-realization as consisting in the development and

exercise of those capacities that make us epistemically and practically responsible agents. In his brief commentary on Green's *Prolegomena*, Brink offers the following description of

Green's view:

[Green] suggests that it is the very capacities that make moral responsibility possible in the first place that determine the proper end of deliberation (§176). Responsible action involves self-consciousness and is expressive of the self. The self is not to be identified with any desire or series or set of desires; moral personality consists in the ability to subject appetites and desires to a process of deliberative endorsement and to form new desires as the result of such deliberations. So the self essentially includes deliberative capacities, and if responsible action expresses the self, it must exercise these deliberative capacities. This explains why Green thinks that the proper aim of deliberation is a life of activities that embody rational or deliberative control of thought and action (§§175, 180, 199, 234, 238-9, 247, 283) (Brink 2003: 40).⁹⁷

This proposed understanding of the nature of the self enjoys some obvious advantages over other accounts we have already considered. Recall that Bradley's pronouncements about the true self were in the end too vague and obscure to be of much help in making his notion of self-realization reasonably precise; this view, by contrast, is quite definite and easily understandable. Also recall Hurka, who held that the good or perfection consists in the development of those properties essential to human nature. We found that there are serious doubts about whether human beings have an essence, and it seems very unlikely that if they do, that anything like rationality is part of it. Since Green is concerned with persons, as opposed to the biological category of human beings, his view is not open to these objections.

One might still wonder why the development and exercise of our capacities for rational deliberation constitutes our good. So far, we have an account of what Green might

⁹⁷ The parenthetical section numbers are Brink's references to Green's *Prolegomena*.

mean by the self, along with the assertion that the good consists in the development of that self. What ties these claims together? It is helpful at this point to recall Hurka's perfectionism from the last chapter. Hurka argues correctly, I think that one of the most attractive features of perfectionism is its assumption that the good consists in the development of our true nature or essence, and self-realization views attempt to exploit this feature. For Green, what is essential to me *qua* person is my capacity for rational deliberation. Thus, there seems to be a straightforward sense in which my capacity for rational deliberation *is* my true self. If we also think that my good consists in the development of my true self, then my good consists in rational deliberation.

How plausible is Green's conception of the good? I shall argue that there are three distinct ways of understanding this identification of the good with rational deliberation, and that none of them is ultimately acceptable. One way of interpreting the view is as holding that our choice or endorsement of a state or activity or way of life after deliberating about the alternatives *makes* that state, activity, or way of life good for the agent; on this view, to say that my good consists in rational deliberation is to say that what is good for me is simply whatever I would choose after the proper sort of deliberation. Due to its (admittedly) vague resemblance to Kantian theories of value, I will for the sake of convenience sometimes refer to this as the Kantian view. One advantage of the Kantian view is that it provides for the possibility of a fairly long and diverse list of goods, since we can imagine choosing any number of things in our deliberations. However, one concern about the view is that the list of goods might be *too* long or diverse. Specifically, it seems implausible that *anything* I happen to choose or endorse upon deliberation would be best or even good for me; the mere fact that I choose a life of pot smoking and tv reality shows after deliberating presumably does not entail

that this is the best life for me (assuming that I have real alternatives).

One might try to solve this problem by insisting on a more robust understanding of proper deliberation. Surely, the possibilities for endorsement after deliberation will depend on the requirements for proper deliberation. If all that is required for proper rational deliberation is the bare consideration of the relevant consequences and alternatives, then pot smoking and reality shows clearly could be chosen: Well, I've thought about it, and the best thing for me is to get high and watch *Fear Factor* as much as humanly possible. Of course, the challenge now becomes that of devising an account of proper deliberation, and it is not at all clear how that would go. One tempting proposal is to say that part of what makes a case of deliberation proper is the extent to which it draws the right conclusions about the good. After all, one of the ways we evaluate our deliberative practices regarding ordinary, garden-variety factual claims is the extent to which these practices result in true beliefs. Why not have a similar requirement for our deliberative practices about the good, so that one of the things that determines whether one's deliberation about one's good is proper is the extent to which it leads to the right conclusions about one's good?

Unfortunately, this conception of proper deliberation is incompatible with the Kantian version of self-realization. For it assumes that there is some fact of the matter about what is good for us which is independent of our deliberative choice, and which our deliberations about the good could be expected to track. However, this is just what the Kantian conception of self-realization denies. Recall that the idea with this account of self-realization is that what makes an activity, state, or way of life good for me is that I choose it after deliberation. Now we are proposing a view which says that there is some independent fact of the matter about what is good, and that the point of deliberation is to track that good.

The first view seems to assume that we confer value on to certain objects through reflective endorsement, while the second assumes a value independent of such endorsement. Since we don't have an account of this proposed independent value—indeed, the whole point of introducing the notion of self-realization was to offer an account of the good—this proposed understanding of proper deliberation and self-realization seems fatally confused.

A more general problem for the version of self-realization that conceives of our reflective choices as conferring value onto the objects of our choices, or as making it the case that certain activities or states are good, is that it's unclear how this view can qualify as a genuinely *perfectionist* account of the good. Recall that subjectivism holds that the good is dependent on the beliefs, desires, commitments, or responses of agents, while objectivism denies this dependence. We saw in an earlier chapter that there are serious problems with subjectivism about the good, and one would expect that most—if not all—of these would be problems for this version of self-realization as well. At the very least, since one of the main points of perfectionist accounts of the good seems to be the rejection of the view that the good is dependent on the beliefs, desires, or responses of agents, it seems safe to assume that this proposal constitutes an abandonment of objectivism about the good, and thus cannot be considered a genuine perfectionist proposal. Moreover, since one of the main motivating factors in the exploration of perfectionist accounts of the good is a dissatisfaction with various forms of subjectivism about the good, a subjectivized perfectionist account would seem to hold little appeal.

A second interpretation of the view that the good or self-realization consists in the development and exercise of our capacities for rational deliberation exploits Green's conception of the self as that part of us that makes us epistemically and practically

responsible agents. It emphasizes the contrast between persons or responsible agents and non-responsible actors, such as children and animals. One important aspect of full-fledged persons is their ability and tendency to regulate their various desires and impulses through the exercise of theoretical and practical reason. For some non-rational actors, such as lower animals and infant humans, there seems to be nothing beyond brute impulses. However, as children start to grow up, they begin to develop and exercise the ability to conceive of their own good as something which is, in principle, distinguishable from, and in potential conflict with, their desires. Unfortunately, for children there remains a strong tendency to be overwhelmed by their desires; in other words, children seem to be especially susceptible to weakness of will. Thus, one of the main tasks in raising children is to train them not to succumb to weakness of will—that is, not to allow their judgments about what they ought to do all things considered to be overwhelmed by their present impulses. Thus, the extent to which we consider children or adolescents to be persons is largely a function of their self-control or resistance to weakness of will. The more that one's actions correspond to one's conception of what one ought to do all things considered, the more one is considered a person.

The view of self-realization we are presently considering is impressed by this understanding of personhood, and concludes that perfection or realization of the self consists in the correspondence of one's actions with one's judgments about what one ought to do, all things considered. That is, self-realization is just the overcoming of, or resistance to, weakness of will. Unfortunately, although this account of self-realization has some *prima facie* plausibility, it is nevertheless untenable. For if all that is required for self-realization is the correspondence between one's desires and one's judgments about what one ought to do,

then one can achieve self-realization simply by altering one's judgments about the good so that they match one's desires. Clearly, if this account is to be at all tenable, it must have something more to say about our judgments about the good and what makes them acceptable or not; mere correspondence between desires and judgments is not enough. However, we are again faced with the fact that the whole point of introducing the notion of self-realization was the hope that it might offer a substantive account of perfection or the good. If we think that the good or perfection consists simply in a match between our desires and judgments about our good, we are left with no direction in forming a conception of the good. This is a fatal objection, since any acceptable account of the good will have to offer some guidance or direction in the forming of our judgments about the good.

There is a third way of understanding the claim that the good or perfection consists in the development and exercise of those capacities that make us epistemically and practically responsible agents. Although this third way is much more plausible than the first two, I shall argue that it nevertheless suffers from its own fatal flaws. Instead of holding that our judgments about the good confer value on certain states or activities, or that self-realization is simply a matter of being resistant to weakness of will, this view holds that what's good for me is, quite literally, the development and exercise of those capacities that are essential to me *qua* person—that is, I achieve self-realization through the exercise of those capacities that make me a responsible agent. It is the development and exercise of my capacities for epistemic and practical deliberation *per se* that constitutes the perfection of my self. Thus, regardless of my present desires, I achieve self-realization through such activities as scientific and other theoretical investigation, and through systematic deliberation regarding what it would be best to do all things considered.

This third reading of Green seems to be the one that Brink endorses, both as an interpretation of Green and as a plausible account of the good. Given our discussion up to now, we can see that a major advantage of this reading of Green is that, unlike its competitors, it does offer clear guidance in developing a substantive account or conception of the good. Although Brink recognizes that Green himself was somewhat ambivalent about the extent to which his version of perfectionism offers practical guidance in developing a conception of the good or of one's ethical duties—for instance, Green suggests both that there is unlikely to be a significant difference in the practical implications of perfectionism and hedonistic utilitarianism and that the voice of individual conscience and the rules of conventional morality are usually adequate guides to conduct (Brink 2003: 73-4)—he nevertheless reads Green as endorsing some rather definitive and substantive positions regarding the content of an adequate perfectionist conception of the good.

We can get a clearer idea of this content by considering some of the implications of this conception of the good. About Green's apparent willingness to settle for an appeal to individual conscience and conventional morality, Brink has the following to say:

This...strand in Green's thought is troubling. It displays a disappointing form of moral complacency that is hard to square with his admiration for the progressive influence of utilitarianism as well as his own calls for liberal reform. Indeed, one would think that Green's own calls for reform—for instance, extending the franchise, establishing state-mandated and state-financed elementary education, and making higher education available to those who are qualified, regardless of economic background—would be plausible examples of the sort of guidance that a perfectionist ethical theory can provide. Presumably, perfectionism also favours poetry over push-pin and soap operas, and meaningful work over drudgery, even well-paid drudgery (2003: 73).

Here, we can see what this version of self-realization would say about our character from a

few paragraphs back who thought about his options and decided that a life of pot-smoking and reality-tv watching would be best. Since such a life would involve, at the very least, the neglect of his powers of rational deliberation and perhaps even the positive undermining of such capacities, this cannot plausibly count as a form of self-realization. The life of our pot-smoking fan of *Fear Factor* is similar to the lives described in Huxley's *Brave New World*, where genetic engineering is used to restrict artificially the capacities of the working classes (Deltas and Epsilons), social and psychological training inhibits them from wanting to develop those capacities they do have, and recreation and drugs (soma) are readily available to keep them content (Brink 2003: 74). The exercise of one's capacities for self-realization, and hence perfection, consists in activities such as going to school, thinking about and debating politics, reading poetry, and engaging in meaningful (and, presumably, intellectually challenging) work.

One difference in the personalities of Green and Bradley was the extent to which they were engaged and interested in real-world political reform; Bradley seems to have been rather indifferent to such issues, whereas Green more closely resembled utilitarians such as John Stuart Mill in his active engagement in political reform. Green's views regarding political reform can help us to flesh out his conception of self-realization. Brink argues correctly, I think that Green's views on political reform suggest a rather robust view about self-realization and what it requires. For Green, a constituent of self-realization is the promotion of the good or perfection of others. Since his political reforms mainly involve policies and programs designed to increase and improve the exercise of capacities for rational deliberation, there is at least *prima facie* evidence that Green sees perfection as consisting in the development and exercise of such capacities. As Brink puts it, Green's perfectionist

liberalism is part of a classical liberal tradition, forcefully articulated by Mill,⁹⁸ which grounds liberal essentials in a conception of the good that prizes the exercise of a person's rational capacities. In Green's version, the good consists of self-realization and the exercise of the very deliberative capacities that make one a moral agent (Brink 2003: 88).

A comparison with Hurka's version of perfectionism would perhaps be helpful here. For Hurka, the good consists in the development, to a high degree, of those properties essential to us as human beings. Theoretical and practical rationality are both essential to human nature on Hurka's view. Thus, perfection consists in the development and exercise of our capacities for theoretical and practical rationality (along with what Hurka calls "physical perfection"). As we saw in the previous chapter, for Hurka theoretical rational perfection is exemplified by the life of the great scientist, whereas practical rational perfection is exemplified by the life of the successful political leader. Green offers conclusions which are similar. Recall that part of what makes us persons is our ability to take a critical stance toward our doxastic impulses, subjecting them to various kinds of scrutiny, and rejecting the ones we have less reason to accept. Since this is an essential part of the self as Green understands it, it seems reasonable to conclude that we can develop our (epistemic) selves by engaging in something like scientific observation and reasoning; thus, Green has a straightforward story to tell about how we can achieve self-realization through scientific inquiry.⁹⁹ Similarly, since we are practically responsible agents in virtue of our ability to

⁹⁸ It's worth noting that Brink interprets Mill as a perfectionist about the good. Although Mill is often taken to be a hedonist, Brink argues that he should be read as holding an objective conception of happiness, so that happiness consists in the development and exercise of certain rational capacities. See Brink 1992.

⁹⁹ Here the contrast with Bradley is instructive. Bradley also thinks that we can achieve self-realization through scientific inquiry along with artistic endeavors but he seems to have no systematic story to tell about how this counts as the realization of the self.

subject our desires to rational scrutiny by doing such things as considering the implications of our actions, weighing alternatives, thinking about the sorts of considerations one might give for or against various courses of action, and so forth, it seems to follow that we can achieve (practical) self-realization through just such activities. Thus, one achieves practical self-realization by leading the kind of life that Socrates counseled – i.e., the examined life.¹⁰⁰

This description of the development and exercise of our capacities for rational deliberation raises a number of difficulties, however. Presumably, it is not the mere exercise of such capacities that we value, but instead the *successful* exercise of them. That is, we presumably don't want to say that practical self-realization consists merely in thinking about the implications of actions, considering and weighing alternatives, and so forth; we want our conclusions about what we ought to do to be the right ones. Unfortunately, it's very difficult to know what to say in the present context about success in practical deliberation. In the case of theoretical deliberation, we have a ready answer: theoretical perfection consists in having the sort of knowledge and understanding of the world which is found in our best scientific theories. Now a moral realist such as Brink might say that we can have theories about the good and the right, just as we can have theories about black holes and inflation, and insist that our deliberation about the good can thus be evaluated in similar terms. However, while that might be true, it does not seem to help us out of our present quandary. For our problem is to specify what is good. Self-realization has been put forth as an account of the good, along with the claim that self-realization consists in the exercise of our capacities for practical

Bradley has to rely on the bare intuition that science (along with art) offers a path of self-realization.

¹⁰⁰ It is worth noting Brink's emphasis on the *process* of rational deliberation:

Pursuit of admissible projects and personal and social relationships that respect persons are both typically more important than the actual realization of projects pursued (1989: 235).

deliberation. But this just seems to suggest that the good consists in deliberation about the good, which seems implausible. We then respond by assuming that practical deliberation about the good, for it to count as genuine self-realization, must lead to the truth about the good. But what is the truth about the good? The truth about the good is self-realization. And what is self-realization? It is the development and exercise of our capacities for rational deliberation. We seem to be arguing in a circle, and a vicious one at that.

This present difficulty with the self-realization view is really just a restatement of the criticism of the first interpretation of self-realization considered several pages back. Such a view seems to assume that there is some account of the good which is independent of the notion of self-realization. However, if there is such an account, then that is what we should be focused on if we are to answer the question, What is good? We have been considering self-realization as an account of the good, and the claim that there is such an account which is independent of self-realization is tantamount to an admission that the notion of self-realization is inadequate as an account of the good. At best, the notion of self-realization offers a link between what we might call good in general with what we might call the personal good. Assume that we have an account of the good in general, and for the sake of illustration assume that the good consists in pleasure, knowledge, and virtue. Self-realization might give us an account of how it is that the pursuit or promotion of these general goods can be part of my personal good. Successful deliberation about the good would require that I recognize pleasure, knowledge, and virtue as goods, which would in turn (presumably) require that I pursue or promote these goods. Thus, the doctrine of self-realization would offer an account of the personal good only in the sense of explaining why it is that an agent should be concerned with the promotion or pursuit of the good in general.

The problem with this somewhat deflated conception of self-realization is that it simply leaves us wondering about the rest of our account of the good. What is the good in general? I used pleasure, knowledge, and virtue as examples for the purposes of illustration, but that choice was really just arbitrary. I could have listed any number of things. What initially attracted us to the notion of self-realization was its promise to specify the good and link the good of the agent with that of others. While it might be up to the later challenge, the deflated version of self-realization fails as a full-fledged account of the good. For it tells us that there is some other story about the good which has little or nothing to do with the notion of self-realization, without giving any hint of where we should turn for this further account.

Given the problems with this deflated version of self-realization, it makes sense to return to what I have described as a more robust version of self-realization — one which holds that the good consists in the development and exercise of our rational capacities. How might this be fleshed out in the case of *practical* rational deliberation? We might think that the good life, on this view, is one where the agent spends a great deal of time and energy thinking about what it is best to do, all things considered. The non-self-realized agent just does whatever he wants, whereas the fully realized agent thinks long and hard before acting, considering all of the relevant implications and weighing all of the relevant available options. The self-realized agent always distinguishes between the intensity and authority of his desires, and often chooses the satisfaction of the less intense desire. Just as a scientist is more of a person in virtue of her devotion to scientific inquiry and reasoning, the practically self-realized agent is more of a person in virtue of the time and energy he devotes to thinking about what would be best, all things considered.

Although such a conception of self-realization offers clear guidance with regard to

how one should lead one's life, it is nevertheless a decidedly implausible and unattractive account of the good. A central problem for this conception of self-realization is that it has no obvious place for subjective goods such as pleasure, happiness, and contentment, and subjective evils such as pain, depression, and anxiety. This is a serious problem since there is reason to think that the good of rational deliberation can, and often does, conflict with these other goods. There is simply no reason to think that the life with the most theoretical or practical rational deliberation will be a particularly happy, pleasurable, or contented life, much less that it will be the *most* pleasurable, contented, or happiest life. Indeed, the exercise of one's capacities for rational deliberation could very well be painful in various ways. For example, when confronted with all of the various implications of one's actions, along with the relevant alternatives open to one, and all of the considerations which might be marshaled for or against a course of action, the agent could simply be overcome with a sense of anxiety or dread. Surely, whatever value there is in spending time thinking seriously and systematically about the origin of black holes or how one ought to act in a given circumstance has to be balanced against the disvalue one might experience in such deliberation. Moreover, there are opportunity costs associated with rational deliberation. Time spent in deliberation is time that could be spent surfing or watching *Seinfeld*; even though I might think that contemplating the origin of black holes is, on balance, better than watching *Seinfeld*, any pleasure I might get from *Seinfeld* has to be taken into consideration, particularly if I will get little or no pleasure from learning about black holes.

Here is an example to illustrate the possible conflict between the good of deliberation and subjective goods. When Joe and Sally go on vacation, say to Disneyland or Yosemite, they employ a division of labor with regard to planning for the trip. Joe does a good deal of

the front-end planning: he decides when they will leave and come back; whether they will fly or drive; if they are driving, whether they will use their own car or a rental; where to get the rental car; where to stay; and so forth. However, the price Sally has to pay for being excused from the burden of such deliberation is that, once she and Joe arrive at the destination, it is up to her to decide their itinerary for that day. She decides the order in which they will explore the park, and Joe trusts her to arrange his day so that he gets as much enjoyment, relaxation, or pleasure as possible. For when Joe goes on vacation, he's like the girl in the Cindy Lauper song: he just want to have fun.¹⁰¹ If Joe has to think long and hard about the best way to explore the park, then unless this deliberation is fun – and I think we have reason to doubt that it will be – then he's missing out on his fun. The last thing Joe want to do is hang around the entrance at Disneyland with T.H. Green deliberating about the best way, all things considered, to tour the park. For there is more to life than the exercise of one's capacity for rational deliberation.

The perfectionist might complain that this example shows that even someone like Joe is committed to the overriding good of deliberation, since his trips typically involve a somewhat complicated division of labor and a set of background judgments about the goodness of various kinds of states and activities. This is of course true, as far as it goes. However, one should note that Joe's approach to taking vacations does not involve anything like the maximizing of the good of deliberation at the cost of other goods. Instead, it involves making a judgment about the value of deliberation relative to other values such as pleasure and enjoyment. Joe judges that the exercise of his capacities for deliberation is only one good

¹⁰¹ I should note that this is compatible with thinking that the fun to be had at Disneyland is significantly, and perhaps qualitatively, different from the fun to be had at Yosemite or the Louvre, and that each can be valuable or desirable in its own way.

among many, and that these various goods can conflict. As best he can, he tries to arrange his life so that he ends up with what seems like an appropriate balance of the various goods. Indeed, I take it that this is what virtually all of us are up to when we engage in practical deliberation. However, this is not the same thing as thinking that the good consists in the development and exercise of our capacities for rational deliberation. Our view assumes that there are goods other than deliberation itself. If a Green-inspired perfectionist wants to agree with this, then we seem to have moved beyond a conception of the good as self-realization, and we are owed an account of this modified view.

Although Brink has, as far as I can tell, little to say about the role of pleasure and related states in Green's perfectionism, Green himself tries to draw some connection between pleasure and self-realization. While most of what Green has to say about pleasure in the *Prolegomena* is in connection with his critique of hedonism, he also has some positive things to say about pleasure. For example, Green tells us that, since there is pleasure in all realisation of capacity, the life in which human capacities should be fully realized would necessarily be a pleasant life (§ 361). He also at least *entertains* the possibility that the perfection of the human soul implies its unimpeded activity, which is pleasure; and that therefore, though in certain stages of the progress towards such perfection there may be for certain persons an abridgment of pleasure, its attainment must be pure enjoyment (§276).

However, it is clear that this latter proposal is not that Green ultimately accepts:

By what right, it may be asked, do we assume that the more developed or perfect state of the human soul is one in which a larger aggregate of pleasure is enjoyed than in the less perfect state? There is pleasure, no doubt, in all satisfaction of desire, there is pleasure in all unimpeded activity. So far therefore as a man has desired the perfection of the human soul, there will be pleasure to him in the consciousness of contributing to that perfection, but not necessarily a greater

amount than he has to forego in order to the contribution. So far as the perfection is attained, again, there will be less impediment to the activity directed to its attainment, and therefore more pleasure in the exercise of the activity. But it would seem at least possible that, according to the plan of the world, the perfection of the human soul may involve the constant presence of a lower nature, consisting in certain tendencies, never indeed dominant, but in conflict with which alone the higher energies of man can emerge. In that case it may very well be that the desire for human perfection, which is the desire for true good, though gradually coming to taste more of the particular pleasure incidental to its satisfaction and to the free play of the action which it moves, as it more fully attains its end, may never be destined to carry men, even in its fullest satisfaction, into a state of pure enjoyment, or into one in which they will be exempt from large demands for the rejection of possible pleasure (§276).

For those of us who insist that an adequate account of the good must include pleasure and related states, Green's claim that the life of self-realization will be a pleasant life is reassuring. However, his admission that the life of self-realization will not necessarily be the most pleasant life is troubling. Indeed, even if it is true that pleasure necessarily accompanies desire-satisfaction or unimpeded activity, it seems that such pleasure could nevertheless be exceedingly slight—for example, the satisfaction of my desire to take my friend's place on the rack would presumably result in very little pleasure. Unless one has very low expectations for the amount of pleasure or happiness needed for a good life, then Green's reassurances about pleasure being a natural by-product of unimpeded activity will not make his perfectionism that much more palatable. Although Green might be right that you can't have self-realization without some pleasure, he gives no reason to think that the amount of pleasure will be sufficient to warrant choosing such a life.

My central complaint against Green's (or we might say, this Green-inspired) conception of self-realization is that it offers a deeply counterintuitive account of the good. It

just seems very implausible that the good life could consist simply in the exercise of our capacities for rational deliberation; a plausible account has to include goods such as pleasure, happiness, and contentment, and evils such as pain, depression, and anxiety. But I want to end by raising a more metaphysical worry about the account as offered. In the previous section, we saw that Bradley's version of self-realization appeals to metaphysical claims which, if not untenable, are at least extremely controversial. An advantage of Brink's reading of Green is that we avoid the appeal to the notion of an interpersonal organic unity, along with many of the other metaphysical and epistemological obscurities of nineteenth-century British Idealism. Instead, we appeal to the notion of a responsible agent, or a person, and hold that the good consists in the development of those characteristics essential to us *qua* persons. Since it is the capacity for rational deliberation that makes us epistemically and practically responsible agents, and thus persons, our good consists in theoretical and practical deliberation; this is how we develop our true selves.

Let's assume for the sake of argument that I am essentially a person, and that the capacity for rational deliberation is essential to that personhood. Even if this is true, to say that I am a person, in the sense of an agent capable of rational deliberation, seems to be a radical under-description of the type of creature I am. For I also have a body of a certain type, and one would think that the good of a rational agent who has the body of a human being would be significantly different from that of a rational agent with some other type of body, or of one with *no* body at all. Green's self-realization account seems to have little or nothing to say about the *human* good, and since we are (perhaps necessarily, though probably not essentially) human beings, this further calls into question the adequacy of Green's account of what constitutes the good for creatures such as us. Indeed, I take it that one of the main

reasons that Hurka, in developing his version of perfectionism, insists on speaking of the essence of *human* nature, even though such a notion is notoriously problematic, is the appeal of offering an account of the human good, and not just an account of one aspect – albeit an important aspect – of our being.

This failure of Green's self-realization view to take seriously our humanness is perhaps at the root of his failure to take seriously goods such as pleasure and evils such as pain. Aristotle famously claimed that human beings are essentially rational, and he also held that this rationality is the distinguishing mark of humanity. Many philosophers have agreed with Aristotle on this score, and have also thought of this rationality as the most important aspect of our nature, and as that which makes us especially important or valuable in the world. Although it is not always clear whether the alleged distinctiveness of our rationality is thought to be what makes it (and us) valuable, what is clear is that many people – philosophers and non-philosophers alike – have seen rationality as especially valuable, and as that which makes human beings more important or more valuable than the other animals. Thus, there is often a tendency for perfectionists to conceive of the good as the development of those properties which, in some sense, set us apart from the animals. We should, of course, be suspicious of any claims about rationality being unique to human beings. More importantly, perhaps, we should be careful not to underestimate the various ways in which we are, in fact, similar to the animals. Our bodies are not *that* different, and we are subject to many of the same goods and evils that they are. If pleasure and contentment are good for my cat, then it would seem that such states should also be good for me. Similarly, if pain and depression are bad for the cat, then such states should be bad for humans. If things can be intrinsically good or bad for dogs, cats, monkeys, cows and chickens, even though they are

not rational agents, then it is hard to see why similar states cannot be intrinsically good or bad for us in a way that is independent of the exercise of our rational agency capacity.

In discussing Brink's reconstruction of Green's understanding of the self and agency, I merely sketched the account without considering any possible problems the view might have. I did this because I wanted to focus on problems with the Green/Brink account of the good in terms of the development and exercise of our capacities for rational deliberation; for I think that even if we assume that the story about personal identity and agency is basically right, the Green/Brink account of the *content* of the good is nevertheless untenable. However, there are, I think, problems with the Brink story about personal identity, and they seem to be related to some things I have just now said about the relation between our personhood and our humanity. My main concern is that I am not convinced that I am *essentially* a person, at least in the way that Brink understands the term. For a person is a responsible agent, and assuming that young children are not responsible agents, and thus not persons, the fact that I can remember things from my early childhood suggests to me that there was a time when *I* was not a person.

Similarly, I think that I can make sense of the thought that when I am very old and on the verge of senility, the things that will happen to my senile self will be happening to *me*. Thus, if I find out that my loved ones have a plan to put me in an institution where I will be treated brutally upon the onset of senility, this is something which will be of great concern to me. For even if I am no longer a rational agent, and thus not able to exercise my rational agency capacity in such a way that my decisions, desires, and actions hook up with those of my earlier life in the way necessary for the persistence of the self, it would seem foolish not to care about the pain in store for me at the institution. Virtually all of us, I take it, if given

the choice between a pleasurable life in the institution and a painful one, would be very much interested in choosing the former. And this does not seem to be because it would be better for *someone*, but because it would be better for *us*. Indeed, the only way we can make sense of our special concern with what will befall us in our dotage is our belief that these are things which will befall *us*.

Perhaps what is at work here is some confusion or conflation between the notions of my self *qua* person and my self *qua* human being. It might be argued that it is only in terms of my self *qua* human being that I could be concerned with what will happen to me in my dotage, or that I could think of the things that happened to me as a very young child. When I think of my self in terms of personhood, then such thoughts make little sense. Maybe our talk about ourselves is simply ambiguous, with it sometimes assuming personhood and other times human-beinghood. If so, we need an account of which way of talking makes the most sense in the context of a discussion of our good. Perhaps the self-realization account which sees the good as rational deliberation is appropriate for us considered as persons, while something more in the direction of hedonism is appropriate for us when considered as human beings. If something like this view is right, it could help to explain why it is so difficult to come up with a plausible, unified account of our good.

This distinction between our selves as persons and ourselves as human beings can also help to illuminate the inadequacy of Brink's account of agency to answer Rawlsian concerns about the separateness of persons. Rawls complains that utilitarianism does not take seriously the separateness of persons, treating individuals as if they were related to the whole of humanity in the same way that an instance or time slice of one man's life is a mere part of his overall life. Brink uses his account of personhood and agency to argue for a limited kind

of corporate agency, according to which the lines between individuals can get blurry if the right sorts of deliberative connections hold between them. What we ought to notice is that, while this might be a plausible view to hold about one's good *qua* person, it is extremely implausible as an account of one's good *qua* human being. For no matter how tight the connections might be between my deliberations and those of my neighbor, there is nevertheless a huge difference between his being subjected to torture and my being tortured. Even in cases where agency is most blurry—say in cases of very close family members—deliberative connections cannot change the fact that the mother's pain affects her in ways that are radically different from the ways in which it affects others in the family. That is why the parent's acceptance of pain in order to protect the children is so obviously a case of self-sacrifice. No amount of psychological connectedness can alter the fact that, for creatures like us, my pain is distinct from your pain; and thus, while we may not necessarily be separate persons, we are nevertheless separate human beings, and a plausible account of our good must recognize that fact.

In this section I've been considering Green's account of self-realization. I wanted to avoid potentially messy exegetical issues, so I focused mainly on Brink's reconstruction of Green's view. I assumed that this is the right interpretation of Green, and went on to discuss its plausibility as an account of the good. I argued that it fails as an account of the good, mainly because it leaves out important subjective goods and evils, such as pleasure and pain. I ended by suggesting that part of the problem with Green's approach is that it ignores an important aspect of our nature. By focusing exclusively on our personhood, or our capacities for rational deliberation, Green's view is unable to account for those aspects of our being which we have in common with other animals. What results is a decidedly

impoverished conception of the good. Indeed, it is not an account of the *human* good at all, and thus is inadequate for beings such as us.

IV. Bradley and Green on Selfishness and Self-Sacrifice

In the previous chapter I suggested that one of the problems with Hurka's version of perfectionism is that it makes use of agent-neutrality in order to account for other-regarding duties. I take it that one of the things that has traditionally made eudaimonistic theories in general, and self-realization accounts in particular, attractive is that they have seemed to offer an alternative account of other-regarding virtue. Instead of holding that practical reason requires one to treat one's own good from a god's-eye point of view, and thus as of no more importance than the good of anyone else, self-realization accounts try to explain how one's own good can be advanced through doing for others. This general approach can be seen in both the Socratic claim that virtue must benefit the agent and the Aristotelian identification of eudaimonia with rational activity in accord with virtue. Bradley and Green both try to account for other-regarding virtue in terms of self-realization; unfortunately, neither is successful and what results is an implausible and distorted conception of practical reason.

Let's start with Bradley. At times, Bradley seems to want to claim that the commonsense contrast between selfishness and self-sacrifice is mistaken or confused in some way (251). Perhaps one must assume this if one is to make sense of the claim that self-realization, i.e., one's own good, can consist largely in the fulfillment of the duties associated with one's station in society. Moreover, however we interpret Bradley's notion of ideal morality, it seems equally unlikely that one's good will necessarily be promoted through the fulfillment of the particular duties associated with this ideal. Although it's easier to imagine

the pursuit of art and science leading to self-realization, even this is doubtful if we take seriously the very real possibility of an extremely unhappy or angst-ridden scientist or artist. Indeed, it is the problem of subjective goods and evils that makes Bradley's conception of self-realization so deeply implausible in the end. Bradley himself, as we have already seen, admits that virtue as commonly conceived can be accompanied by a good deal of unhappiness as commonly conceived. Moreover, Bradley tells us that although we cannot hate the good *qua* good, we can hate it for what accompanies it: Certainly in one way I may hate good. I may loathe it, because, though I desire it, it brings me perpetual pain and weariness (306n1). Of course, we have also already seen that Bradley admits pleasure as a good, which seems to commit him to the conclusion that the good of pleasure can conflict with that of self-realization, and to the conclusion that there is more to the good than mere self-realization as Bradley understands it.

We have already seen examples of tension in Bradley's thinking about the relation between self-realization and selfishness. Although Bradley wants to explain, and presumably justify, other-regarding morality by tying it to the good of the agent, he also identifies selfishness as 'thinking only of yourself' and offers the following explanation:

Thinking only of oneself implies first that we think, that we are self-conscious reflecting beings; and hence it seems a misnomer to call a beast or a young child selfish. Secondly, we think of nothing but ourselves; and this means that the ends we set before us have not an objective content which is desired for itself, and without regard to its relation to our private selves. The selfish man, so far as he is selfish, has objects of desire which are not subordinated to any principle higher than his private satisfaction. If you ask what is the general end which includes his ends, you can point to none; but you find he treats all objects as means, that he cares for none in itself, but will sacrifice any with readiness; and when you inquire what is common to them all, you find that they minister to his personal comfort; this comfort being a certain

quantum of the pleasant and of absence of pain, which satisfies him, and which he either consciously aims at or unconsciously uses as a measure of all objects of desire (274-75).

Since Bradley admits (at least at times) that pleasure is a good, this identification of selfishness with the pursuit of one's private pleasure shows that there can be a gap between one's good and one's self-realization, as Bradley understands it; that is, the pursuit of one's good can be genuinely selfish, and self-realization can require the sacrifice of at least some aspect of the agent's good.

In a note added for the second edition of *Ethical Studies*, Bradley includes a discussion of selfishness which seems to recognize the way in which one's good and one's duty can diverge. Bradley seems to have come to accept the more commonsense understanding of selfishness as consisting of an exaggerated or undue preference for one's private good at the expense of the good of others:

Self-realization may be immoral because selfish, if, and so far as, it means undue subordination of others' welfare to one's own, in however high a sense the latter is taken. If the devotion of oneself to a pursuit involves the sacrifice of others' welfare beyond what is justifiable in the case, that is selfish...It does mean realizing oneself in a certain way without due regard to others' welfare (276n1).

This seems to be a fairly radical departure from the program which seeks to explain the importance or authority of other-regarding reasons or virtues in terms of the good of the agent. The idea now seems to be that one is supposed to balance one's own good with the good of others, and if one gives too much weight to one's own good in relation to the good of others, then one is being selfish. As far as I can tell, this seems to fit very closely commonsense thinking about the nature of selfishness. The mere fact that I am pursuing my own good is not sufficient to make me selfish. However, if I would prefer the torture and

murder of someone who I shall never meet to a mere inconvenience to myself, such as a slight scratching of my finger, then that would be a paradigmatic case of selfishness.

This notion of selfishness as consisting in an exaggerated or unwarranted favoring of one's good at the expense of the good of others is perhaps forced on Bradley by the recognition that self-realization, as he understands it, can require a genuine sacrifice of the agent's own good. Interestingly, the recognition of pain and pleasure and related states as genuine goods and evils can help to make some of Bradley's other claims about self-realization more plausible. I have been taking Bradley to task for failing to include subjective goods and evils into his official account of self-realization, and have made a great deal of hay over the apparently undeniable fact that self-realization, as Bradley understands it, can be painful or lead to less happiness for the agent. This might lead one to conclude that the sorts of things Bradley points to as constituting self-realization cannot possibly benefit the agent. However, if we include the subjective goods and evils and recognize that self-realization can mean less of such goods and more of such evils, then the claim that self-realization is good for the agent becomes less implausible. On this view, something like the fulfillment of one's station in society does benefit the agent insofar as it is a constituent of self-realization. However, it is only a *prima facie* or *pro tanto* good, and it can be outweighed by other goods, such as happiness or the absence of pain. On this view, practical rationality requires the balancing of the good of self-realization with other goods. Although one might still doubt that Bradley has made the case that fulfilling one's station in society is really part of one's self-realization or good—perhaps because one doubts Bradley's account of our social essence—the admission of a plurality of goods makes the implications more palatable.

Since Green doesn't share Bradley's views about the *content* of self-realization, any

problems with Green's account of selfishness and self-sacrifice will be of a different nature. How successful is Green at tying the good of the agent to the promotion of the good of others? If we stick with Brink's reconstruction of Green, we can see that while Green can justify a certain degree of other-concern via the good of the agent, there are nevertheless some significant gaps. Recall that, on the Green/Brink view, it is through a certain kind of deliberative connectedness that the good of others becomes part of my good. One problem has to do with the scope of one's concern. The Green/Brink view as described seems to suggest a reason to be concerned with the good of those I am already closely connected to e.g., family members and close friends but it gives no obvious reason to concern oneself with the good of people one is not already connected to. Indeed, it is clear that Green and Brink both want to extend the interests of the agent to cover all of humanity. But wouldn't this just reintroduce all of the problems associated with agent-neutrality? Brink tries to solve this problem by appealing to Broad's notion of self-referential altruism, according to which the agent's concern has a universal *scope*, but a variable *weight* (Brink 2003: 52-59). According to Brink, self-referential altruism is impartial in one sense in so far as it recognizes non-derivative reason to benefit others but it is partial in another sense in so far as the weight or strength of the agent's reasons are a function of the relationship in which she stands to potential beneficiaries (2003: 56).

Brink's self-referential altruism results in a view of proper moral concern which can be thought of in terms of a set of concentric circles. At the center is the agent, and she gives the most weight to herself, since this is where we find the greatest degree of deliberative connectedness. Next we find family members and close friends, followed by colleagues or business associates, fellow citizens, and so forth, eventually extending to those others with

whom we have little or no connection. The agent gives greater weight to the good of those who are in the inner circles – e.g., family members are given greater weight than fellow citizens. Although the weight given to the good of various other agents is variable, one's moral concern must have a universal scope, so that it includes even Aristotle's remotest Mysian.

While it is understandable that Brink and Green want to account for a universal scope for moral concern, the appeal to self-referential altruism seems ill-equipped for the task. *Ex hypothesi* we have no connection with the remotest Mysian, so it is unclear how his good can be part of our own good. Brink concedes that this is a potential problem for his view, but he offers the following reply:

If the remotest Mysian and I stand in no relations of psychological connection, then his good is not already part of mine. So I can have no backward-looking eudaimonist reason to be concerned about him. But I can have forward-looking reasons. For it is now within my power to interact with him, and all the reasons for cultivating interpersonal self-extension apply and provide a forward-looking rationale for concern. Even when the remotest Mysian and I have no prospect of further interaction, my assistance will enable or facilitate his pursuit of his own projects, and this will make his subsequent actions and mental states dependent upon my assistance. Indeed, other things being equal, the greater the assistance I provide the greater is my involvement in his life. To the extent that another's actions and mental states are dependent upon my assistance, I can view the assistance as making his good a part of my own. Assistance to the remotest Mysian earns me a share, however small, of his happiness, much as care and nurture of my children ground posthumous interests I have in their continued well-being. This is why Green thinks that self-realization involves contributing to a larger, more permanent, and comprehensive good. If so, it explains how a eudaimonist can legitimately seek a universal common good, of the sort Green contemplates (2003: 55).

The main problem with this line of thinking is that it assumes that the agent's good consists in

the extension of her deliberative capacities. It describes how one can get involved in the life of the remotest Mysian and thereby further influence the world, but ignores the fact that the payoff for such involvement is bound to be disappointing to someone not already convinced of the value of this sort of increased deliberative control and involvement. In particular, if one includes in the cost and benefit analysis the gains and losses of subjective goods and evils brought about by getting involved with the remotest Mysian, then it will be much less apparent why it is in one's interest to do so.

But there is a more fundamental problem with this approach to explaining why we have reason to be concerned with the good of others. It seems to require that we care about the good of others for purely instrumental reasons. I am required to care about everyone from my spouse and children to the remotest Mysian simply because this is the way for me to promote my own good. While this *might* be an acceptable account of my concern for the remotest Mysian, it seems decidedly inappropriate as a way of thinking about the good of those near and dear. Surely, I am to care about family and friends for their own sake, not because doing so will make me better off in the long run. Even if we think that the agent's good can conflict with the good of family and close friends, and even if we think that it sometimes makes sense for the agent to sacrifice herself for the good of others, it seems a profoundly distorted conception of practical reason which would explain such reasons in terms of the agent's self-interest. Moreover, even in a Good Samaritan-type scenario where I come across the remotest Mysian suffering on the side of the road and I can relieve his suffering with very little cost to myself, to attempt to explain the demand that I help him in terms of self-interest just seems to be the wrong way to think about it. The Mysian's suffering, along with my ability to help at little or no cost, seems to give me reason to help in

and of itself.

One way of dealing with these various problems is to adopt not only a pluralistic account of the good, but also a pluralistic account of practical reason. Such a view could recognize prudential, (other-regarding) agent-relative, and (other-regarding) agent-neutral reasons for action. Just as there is no one fundamental type of good to which all other goods can be reduced, there is no aspect of practical reason which has sole, ultimate authority. This view would recognize that the suffering of the remotest Mysian can act as an agent-neutral reason for me to ease his suffering. However, this agent-neutral reason can conflict with both other-regarding agent-relative reasons such as special duties I might have to my family or friends, and with prudential reasons. Just as determining what is in my self interest requires making some kind of judgment about what is best for me all things considered, where this involves choosing among potentially conflicting goods, determining what I ought to do all things considered requires that I make some sort of judgment about the relative weight or importance of potentially conflicting reasons for action. Bradley's *second* thought about the nature of selfishness seems to come closest to capturing this pluralistic conception of practical reason. The mere fact that action A would best promote my self interest is not necessarily a decisive reason to do A. If doing A would cost someone else, even an absolute stranger, dearly enough, and the benefit to me of doing A is slight enough, then it would be selfish for me to do A, so selfish perhaps that it just does not make sense, all things considered, for me to do A. Of course, it will be a substantive question what counts as enough of a burden on others to outweigh my prudential reasons, but at least this view offers an honest and plausible account of how other-regarding reasons relate to reasons of self-interest.

Conclusion

In this chapter I've considered a version of perfectionism known as self-realization. According to self-realization accounts, the good for an agent consists in the development of the self. Moreover, self-realization theorists typically attempt to link the good of others to the good of the agent, so that the agent can achieve self-realization through promoting the good of others. I began by discussing Bradley's version of self-realization, arguing that, although his view has certain attractive implications, it is nevertheless an untenable hodgepodge of otherwise intuitively attractive views about ethics, with no clear notion of self-realization to tie them all together. For Bradley self-realization consists in the fulfillment of the duties associated with one's station in society, along with another somewhat obscure set of duties which he calls 'ideal morality,' and the pursuit of truth and beauty. He also suggests that pleasure is a good as well, though he emphasizes that it is not *the* good, thus rejecting hedonism. While this may be an attractively diverse conception of the good, Bradley never develops a clear and plausible conception of self-realization which could unify these various aspects of morality. Thus, Bradley serves as a useful case study in the limits of self-realization theories to account for our considered ethical judgments, and offers a clear illustration of what one might call the push toward pluralism.

A more fundamental problem with Bradley's view, perhaps, is its reliance on the notion of an interpersonal organic unity. Although Green is usually seen as holding basically the same metaphysical and epistemological views as Bradley, David Brink suggests a reconstruction of Green's view which does not rely on the notion of an interpersonal organic unity in order to explain how the good of others can be a constituent of the agent's good. I assumed that Brink's interpretation of Green is basically correct, and went on to consider its

plausibility and attractiveness as an account of the good. I argued that the Green/Brink account of the good in terms of the development and exercise of our capacities for rational deliberation is ultimately untenable, mainly because it leaves out important subjective goods such as pleasure, happiness, and contentment, and evils such as pain, depression, and anxiety. Although these capacities may be extremely important since they are what make us epistemically and practically responsible agents, and thus persons, there must be more to a good life than the exercise of such capacities. A second major problem with self-realization accounts is that they ultimately depend on, and encourage, a profoundly distorted conception of practical reason, along with an implausible account of the relation between selfishness and self-sacrifice.

Thus, we reach a conclusion here which largely mirrors one we reached in the discussion of Thomas Hurka's version of perfectionism in the last chapter. Hurka admits that there is no room for pleasure and pain (and related states) in pure perfectionism, and argues that, if we insist on including them in our conception of the good, then we must resort to a pluralistic account of the good. This seems to be the situation we find ourselves in at the end of our consideration of self-realization versions of perfectionism. We could, of course, turn to some other version of perfectionism to see if it could account for the diversity of goods. However, it seems unlikely that such a strategy would be fruitful. Instead, in the next chapter I shall develop a pluralistic account of the good, explain what such a pluralism amounts to, and consider some of its implications for ethical theory and practice.

Chapter 5 Value Pluralism and Ethical Judgment

In previous chapters, I criticized various accounts of the good, including hedonism, the desire-satisfaction view, and two different versions of perfectionism. I argued that none of these accounts of the good is plausible. Although hedonism and perfectionism both seem to have part of the truth about the good, ultimately they each leave out important goods: hedonism leaves out goods such as knowledge and achievement, while perfectionism ignores such goods as pleasure and happiness.¹⁰² Thus, neither hedonism nor perfectionism offers a complete recipe for a good life; any life guided by either hedonism or perfectionism would be lacking in important goods.

Although the desire-satisfaction view avoids the incompleteness problem which plagues these other views, it is nevertheless untenable as well. A simple desire-satisfaction view which identifies one's good with the satisfaction of whatever desires one actually has cannot possibly be right. All too often we form desires the satisfaction of which turns out to be harmful. This commonplace phenomenon has led philosophers such as Brandt and Railton to identify the good with the desires one would have under suitably idealized conditions. In chapter 2, I argued that even these more sophisticated versions of the desire-satisfaction view have implications so counterintuitive as to render such a position ultimately untenable. More fundamentally, the desire-satisfaction view is false to the phenomenology of value: while it

¹⁰² Here I assume a modern, subjective conception of happiness of the sort discussed by Kraut in *Two Conceptions of Happiness*. Of course, some philosophers—particularly those influenced by Aristotle—argue for an objective conception of happiness, according to which happiness is constituted by something like virtuous or rational activity. I think that such views are best described as accounts of human flourishing or the human good, and that, to avoid confusion, we ought to restrict the term *happiness* to its more modern, subjective sense.

assumes that things are made valuable by their being desired under certain circumstances, our actual experience of judgments of value suggests that we value certain things because they are valuable or good. We desire pleasure, for instance, because it is good, and are averse to pain because it is bad; they are not *made* good or evil by our desires or aversions. If I am in severe pain, my life is, to that extent, going poorly, and I cannot eliminate the evil merely by forming a desire for the pain.

In this chapter I put forth my own account of the good. I begin by offering what I take to be strong candidates for constituents or components of the good, and then consider the extent to which these goods can be unified in some way. I argue that pleasure and freedom from pain (along with related experiential states), proper emotional responsiveness, and knowledge or understanding all deserve a place on our list of goods. Although this list is fairly short, it is also *diverse*: these goods are distinct from one another, and there does not seem to be any underlying principle or value tying them all together. Thus, we end up with pluralism about the good. But some forms of pluralism are more promiscuous than others, and we might wonder if we can nevertheless achieve *some* degree of unity or explanatory coherence even within a pluralistic framework. I argue that we can. The good of an individual depends on the kind of being that it is.¹⁰³ Human beings are of a complex nature;

¹⁰³ It is important to note here that I am not claiming that the good of an individual has to be relativized to the kind of being it is in such a way that cross-species judgments of quality of life become impossible or nonsensical. I take it that we are committed to the judgment that the life of a normal, fully developed human being is of a higher quality than that of a normal, fully developed cat; i.e., it is better to be a human than a cat. Otherwise, it is difficult to see how we could make sense of our considered judgment that the loss of one's rationality constitutes a significant reduction in one's quality of life. Thus, if we are confronted by aliens who share our rational agency capacity but completely lack our capacities for interpersonal sympathy, it is an open question whether the lack of such capacities makes the aliens' lives worse.

although we share a great deal with the rest of the animal kingdom, our capacities for rational deliberation make us quite different in important ways. In short, we are embodied rational agents, and as such our good will be correspondingly complex, consisting both in the having and avoidance of certain kinds of feelings, emotions, and sensations, and in the development and exercise of our capacities for rational deliberation.¹⁰⁴ The sorts of monism considered in previous chapters make the mistake of ignoring the complexity of our nature, and thus end up with accounts of the good ill suited to the sorts of beings we are. In short, they achieve theoretical unity and explanatory coherence at the expense of plausibility and livability, resulting in accounts of the human good which inspire, at best, mere lip service.

Although the appeal to the complex nature of human beings offers some explanatory unity to our judgments about the good, we should not lose sight of the fact that this view is still a form of pluralism. Although it explains why, for instance, pleasure and knowledge are both good, it offers no obvious guidance regarding tradeoffs between these two goods. Cases where the goods conflict will require the use of one's judgment to determine whether, for instance, this amount of knowledge or understanding is worth sacrificing that amount of pleasure or contentment. Moreover, as we shall see, what counts as proper emotional responsiveness will often be a matter of judgment, and while we might be certain that anger or sadness is the appropriate response to a given situation, it will not be obvious *how much* anger or sadness should be felt or exhibited. This appeal to the use of judgment in determining how competing goods ought to be balanced against one another or what counts as

¹⁰⁴ I should emphasize that I am not claiming that such properties or characteristics are *essential* to human nature; it is simply an important fact about human beings that we have such capacities. There are strong reasons for doubting that human beings have an essence of any sort, and the claim that knowledge or rationality is essential to human nature is especially implausible. See my earlier discussion of Hurka's claims about the human essence.

the appropriate emotional response in a given situation is admittedly problematic. In fact, some might insist that what we need is a theory of proper judgment which will give us determinate answers in such cases. However, I think that this impulse toward determinacy, though understandable, is what leads philosophers to embrace the sorts of specious monistic accounts previously considered. If we are to have an account of the human good which we can take seriously, then we will have to make some sacrifices with regard to theoretical unity, simplicity, and determinacy.

I: The Good Things in Life: Introductory Remarks

What makes a life go well?¹⁰⁵ In his review of philosophical thinking regarding the good, William Frankena¹⁰⁶ offers the following rather lengthy list of typical candidates for objects of intrinsic value:

Life, consciousness and activity
 Health and strength
 Pleasures and satisfactions of all or certain kinds
 Happiness, beatitude, contentment, etc.
 Truth
 Knowledge and true opinion of various kinds, understanding, wisdom
 Beauty, harmony, proportion in objects contemplated
 Aesthetic experience
 Morally good dispositions or virtues
 Mutual affection, love, friendship, cooperation
 Just distribution of goods and evils
 Harmony and proportion in one's own life
 Power and experience of achievement
 Self-expression
 Freedom
 Peace, security
 Adventure and novelty

¹⁰⁵ This way of putting the issue is borrowed from Derek Parfit. See *Reasons and Persons* (493-502).

¹⁰⁶ Frankena 1973 (87-88).

Good reputation, honor, esteem, etc.

Although this list has some degree of structure, with, for instance, certain goods such as knowledge and understanding being grouped together, the overall impression is of a motley collection of goods, with little or nothing linking them together or organizing them into a coherent whole. Another thing to notice about this list is that certain items cannot serve as constituents of the personal good since they are not properties of a given life. Whereas things such as health, strength, pleasure, contentment, knowledge, and good reputation are possessed or exemplified by individual persons, truth is a property of statements or beliefs while the just distribution of goods and evils is a state of affairs. Neither of these items can, in and of itself, make an individual's life go better. This is not to say, however, that pursuit or achievement of such non-personal goods cannot be constitutive of one's personal good. Perhaps in the end we will want to say that one's life is made better by being properly oriented toward non-personal goods such as truth, distributive justice, a healthy ecosystem, and so forth.¹⁰⁷ But for now, we should bear in mind the distinction between personal and non-personal goods.

Frankena notes that certain philosophers offer lists which are considerably shorter than the one he offers. He also notes various attempts to organize the various goods into categories such as biological, physical, mental, social, and spiritual. Such schemes of categorization are reflected in recent work in the Natural Law tradition, with John Finnis for example listing life, knowledge, play, aesthetic experience, sociability or friendship, practical reasonableness, and religion as fundamental goods.¹⁰⁸ Mark Murphy offers a very

¹⁰⁷ Hurka offers an account of virtue along these lines in *Virtue, Vice, and Value*. He takes virtue to consist in loving or being properly oriented toward the good, and proposes pleasure, knowledge, and virtue as a provisional account of the good. Obviously, many other things could be added to the list of objective, non-personal goods.

¹⁰⁸ Finnis 1980 (85-90).

similar account, adding excellence in agency, inner peace, and happiness to his list.¹⁰⁹ Both authors tie these various goods to certain capacities or needs human beings are taken to have. We can find shorter lists outside of the Natural Law tradition as well. For example, W.D. Ross famously offers pleasure, knowledge, virtue, and the just apportionment of pleasure to desert as fundamental goods.¹¹⁰ While Ross doesn't think that there is any further feature underlying these goods or tying them together (other than the non-natural property goodness), he does try to limit the list of goods to these four, arguing that other putative goods such as friendship and aesthetic enjoyment are ultimately reducible to one or more of his four fundamental goods.

My central aim in this chapter is to consider the extent to which we can develop a plausible list of goods which is shorter and more structured than the one offered by Frankena. While Frankena's list is longer and less structured than need be, we shall nevertheless see that there are several fundamental goods which any plausible theory must recognize. Since each of these goods is fundamental, and they can and often do conflict with one another, our view is decidedly pluralistic. Let's now turn to the issue of which things make a life go well.

II. Pleasure and Freedom from Pain

Although hedonism, the view that pleasure is the *only* intrinsic good and pain is the *only* intrinsic evil, seems deeply implausible as a complete account of the human good, I think the intrinsic goodness of pleasure and the intrinsic badness of pain are undeniable.¹¹¹ Indeed,

¹⁰⁹ Murphy 2001 (100-135).

¹¹⁰ Ross 1930 (134-141).

¹¹¹ Surprisingly, a fairly large number of philosophers do in fact deny that pleasure is intrinsically good. See, e.g., Richard Kraut, "Desire and the Human Good." The Ross of *The*

there seems to be a number of mental or experiential states which are either good or bad purely in virtue of their intrinsic qualities. Examples of intrinsically valuable experiential states include sensuous pleasures, such as that derived from eating tasty food or engaging in certain sexual activities, and intellectual pleasures such as that which might be derived from successfully working a complex mathematical proof or coming to understand some complex scientific theory. The widespread recognition of the value of pleasure and freedom from pain is no doubt one of the reasons that hedonism was one of the first systematic attempts to specify the good, and why non-hedonists so often feel a need to explain away the apparent intrinsic value of pleasure, for example by admitting that it is good, but then arguing that it is only instrumentally, or perhaps conditionally, good.

A tremendous number of our everyday judgments about both morality and self-interest support the claim that pleasure is intrinsically good and pain is intrinsically evil. For example, a paradigmatic case of moral obligation is the duty to refrain from inflicting physical pain on another human being. Moreover, actions which unnecessarily cause various kinds of mental anguish are also typically thought to be immoral. Similarly, paradigmatic acts of beneficence include those which involve the spreading of pleasure, happiness, and contentment and the reduction or alleviation of pain, distress, and sadness. The good of animals seems to be largely, if not exclusively, a matter of the kinds of mental states they experience; it is hard to imagine evaluating the quality of life of one's dog without referring to the sorts of pleasures and pains he is experiencing. We also typically think of self-

Right and the Good thought that pleasure was intrinsically good; however, he changed his view by the time of the publication of *The Foundations of Ethics* nine years later. In

Pleasure and Reflection in Ross's Intuitionism, Philip Stratton-Lake discusses Ross's treatment of pleasure in the two works and is himself quite sympathetic with the later Ross's rejection of the intrinsic value of pleasure.

interested action in terms of the pursuit of pleasure and avoidance of pain; indeed, the paradigmatic selfish person is one who has a disproportionate regard for his own pleasure and pain. One of the reasons we resent such a selfish person is our sense that he is getting something good at the expense of others. It is hard to see how we could make sense of such commitments absent the belief that pleasure and freedom from pain are intrinsically good, at least for the those who experiences such states.

Moreover, in explaining the behavior of people we constantly find ourselves referring to the particular mental states they are taken to have. This is true with regard both to *motivational* explanations and *justificatory* explanations. If I go to the dentist for a filling and ask the dentist for novocaine, she does not give me a puzzled look, as if she had no idea why anyone would want to mask the pain from the drill; indeed, if she did express such puzzlement, I would probably look for a new dentist. If someone says that he spends his time reading Dickens instead of Trollope because he gets more pleasure or enjoyment out of Dickens, we take this as at least *prima facie* a good reason for him to spend his time that way; if we think that he should nevertheless be reading Trollope instead, we owe him and ourselves for that matter some reason or argument to that effect. If, by contrast, someone says that he prefers Dickens to Trollope because he gets more pleasure or enjoyment out of *Trollope*, then one cannot help but be puzzled. To make sense of this, we need a further explanation. For instance, we need to hear that this fellow believes that, as a general rule of thumb, the more pleasant or enjoyable writer is usually not the more intellectually or morally edifying and that it is the development of one's intellect and character which he thinks is more valuable; or perhaps we could be told that he is filled with guilt and self-loathing, and thus thinks of himself as unworthy of pleasure or enjoyment, and therefore avoids it when he

can. The justificatory power of such explanations will depend on where we stand on the relative importance of edification versus pleasure, or on whether or not the reader is deserving of pleasure. But these explanations do not suggest that pleasure lacks intrinsic value; they merely suggest that there are things other than pleasure which can contribute to one's good, or they suggest that one can have reason to pursue something other than one's good.

Although it seems obvious and undeniable that pleasure is an intrinsic good, it can be surprisingly difficult to say precisely what pleasure is. So far, I have glossed over this potential problem by using the terms *pleasure* and *pain* fairly loosely. However, it is important to note that there are a number of concepts which are related to, but in some sense distinguishable from, pleasure, such as happiness, contentment, satisfaction, enjoyment, joy, delight, cheerfulness, gladness, elation, ecstasy, and beatitude (to name just a few). Although some of these terms can sometimes be substituted for one another without a loss of meaning e.g., *pleasure* and *enjoyment* are often used, by philosophers at least, pretty much interchangeably the connotations of some of these terms diverge. For example, if Bob dresses up as Santa Clause and distributes presents to the neighborhood children, it is natural to say that he is spreading the joy and cheer of the Christmas season. However, if Bob goes back to his job as a prostitute after the holidays, it would be linguistically odd to describe his giving out of sexual favors as further instances of spreading joy and cheer. Indeed, if one did say such a thing, it would most naturally be interpreted as a joke. It is a joke because Bob's customers are most likely to be in the market for (and are more likely in fact to get) sensual pleasure, which is significantly different from emotional states such as joy, cheer, and

delight.¹¹²

The linguistic oddness of describing Bob's customers as deriving joy, delight, and cheer from their dealings with him *qua* prostitute might lead us to conclude that what separates joy, delight, and cheer from pleasure is that the first three have to do with emotional states, while pleasure refers to bodily sensations. However, while it is true that joy, delight, cheer, gladness, and contentment do seem to be emotional states, it is not the case that pleasure is restricted to bodily or sensual pleasures. We have already seen that we can sensibly talk of the pleasure one gets from reading Dickens or Trollope, and one can also talk about the pleasure one gets from working a mathematical proof, or contemplating the metaphysical status of time, or reflecting on one's past accomplishments, or anticipating the success of a loved-one, or watching children run in the grass, and so on. This is the way in which pleasure can be used more or less interchangeably with enjoyment. Both are quite broad terms, and while they both refer to mental states, these can include a wide range of emotions, feelings, and sensations.

One might also think that there is an important distinction to be made between pleasure and happiness along the lines of the latter being a dispositional property or state and the former a transient sensation. It is certainly the case that happiness is sometimes used in this way. We can say of someone that she is happy even though she is asleep or unconscious; moreover, we can make perfect sense of the statement, I don't know why she is crying; she is such a happy person, so long as we understand happiness in dispositional terms. Similar things can be said about the relation between cheerfulness and transient feelings of cheer,

¹¹² I do not mean to imply that it is *impossible* for one to experience such emotional states via intercourse with a prostitute. It is just that they are extremely rare given the impersonal nature of such exchanges.

where the former is thought of as consisting in a disposition to experience transient feelings of cheer in certain circumstances. However, while pleasure certainly does not seem to admit of anything like a dispositional analysis, it is simply not the case that happiness *must be* understood in dispositional terms. We can make perfect sense of the idea of happiness as a transient state, so that one can sensibly say the following: I've never seen Joe so happy as when he found out he had won the lottery. Similar things can be said about cheerfulness: it does not have to be thought of as a dispositional property; we can say that Joe is cheerful i.e., he exemplifies cheerfulness in the sense that he is in a rather cheerful mood due to the Christmas present he just received from Bob.

One could continue in this ordinary-language vein for quite some time if one were inclined, trying to tease out the various conceptual relations holding among these various notions; however, my sense is that little would be achieved. Some of these terms refer mainly, or perhaps even exclusively, to certain emotional states. But pleasure and enjoyment and even happiness seem to be quite elastic, and the most we can say about them is that they refer to mental states which we take to be intrinsically desirable. Similar things can be said about states such as pain, anxiety, dismay, depression, agitation, anger, distress, sadness, and fear. Some of these concepts for example fear and anger are quite thick, having a rather specific meaning. But others for example pain and distress seem quite elastic, and can refer to a wide range of negative mental states in a way that largely parallels the role of terms such as pleasure and enjoyment. However, in spite of the vagueness and ambiguity of these various terms, we can still say that they all refer to mental states which we take to be intrinsically undesirable.

As long as we keep in mind the elasticity of the terms pleasure and pain, and do

not assume, for example, that they only apply to sensations, we can define pleasure as any intrinsically desirable experiential state and pain as any intrinsically undesirable experiential state, where the desirability or undesirability is purely a function of the phenomenological qualities of the experiential state. There are certain mental states – feelings, emotional states, and sensations – which are intrinsically desirable, and others which are intrinsically undesirable; they are good or bad, desirable or undesirable, in virtue of their internal character, or we might say, in virtue of how they feel. A couple of examples might help to clarify what I mean by this. Go back to our friend who prefers Dickens over Trollope because Trollope gives more pleasure. We speculated that he might have this preference for Trollope because of his belief that the more pleasant read is the morally or intellectually less edifying, along with his belief that the development of character and intellect is what is most valuable. There are at least two ways in which this could be taken. First, it could be meant that the feelings or other mental states associated with the exercise of good moral character or a highly developed intellect are superior *qua* experiences to other kinds of pleasurable experiences, say the sorts of feelings or emotional states one gets from inferior art. If this is what is meant, then I take it that we are just saying that our friend gets more enjoyment from the exercise of character and intellect than he does from consuming inferior art. If, on the other hand, it is meant that our friend values the development and exercise of intellect and character for reasons that are independent of the way such activities feel – perhaps because he takes them to have a certain kind of perfectionist value – then his judgments and preferences are tracking something other than the internal nature of the experience.

One of the items on Frankena's list, beatitude, offers another good illustration of the distinction between states that are valued purely for how they feel and those which are valued

at least partially for other reasons. Beatitude is typically thought of as supreme blessedness. It is the greatest good that man can achieve and usually is thought to require standing in a certain relation with God. What is it precisely that makes beatitude the greatest good for a human being? There are at least two things one might say. First, one might say that beatitude is the greatest good simply because of the experience itself. The idea would be that no experience offers the sort of joy, contentment, fulfillment, and satisfaction that beatitude does, and this is why it is God's greatest gift to humanity. A second possibility is to say that it is the fact that the blessed person stands in a certain relation with God which makes it the greatest good. On this second interpretation, beatitude presumably has a certain value *qua* experience, but it contains additional value for the blessed due to its involving a certain relation with, or orientation toward, God. And of course these two views could be combined such that beatitude is thought to involve the greatest subjective experiences one could possibly experience and to have additional value in virtue of the relation in which the blessed stands to God.

When I say that certain experiential states are valuable or desirable for their internal character, or simply for how they feel, what I have in mind is the value beatitude would have for an agent independently of any accompanying relation with God, or the value of the experience of exercising one's aesthetic judgment or intellect or character independently of any perfectionist value such activities might have. This account of pleasure is not unique in the history of philosophy. Its most notable proponent is Henry Sidgwick:

...for my own part, when I reflect on the notion of pleasure, using the term in the comprehensive sense I have adopted, to include the most refined and subtle intellectual and emotional gratifications, no less than the coarser and more definite sensual enjoyments, the only common quality that I can find in the feelings so designated seems to be that relation to

desire and volition expressed by the general term desirable ...I propose therefore to define Pleasure ...as a feeling which, when experienced by intelligent beings, is at least implicitly apprehended as desirable or in cases of comparison preferable (127).

This account of pleasure in terms of intrinsic desirability is controversial. However, when we consider the prevailing alternative accounts of pleasure and pain, we see that the intrinsic desirability account is the most plausible of the options. One alternative account of pleasure is what we might call the *Distinctive Feeling* view, according to which pleasure is a certain kind of introspectively identifiable feeling analogous to sweetness or butterflies in the stomach.¹¹³ Bentham probably thought of pleasure and pain in such terms; certainly, his distinguishing among pleasures and pains in terms of intensity, duration, certainty or uncertainty, propinquity or remoteness, fecundity, and purity seems to presuppose a conception of pleasure as a unitary sort of experience, and suggests that particular instances of pleasure and pain can differ only along quantitative lines.¹¹⁴ However, as we have seen, such an account of pleasure is simply not plausible given the considerable phenomenological diversity among different types of pleasure. Consequently, the idea that pleasure is a unitary experience which attends certain sorts of activities or events is a view which very few philosophers take seriously.

C.D. Broad argues for an account of pleasure which is quite similar to the *Distinctive Feeling* view, but which is designed to account for the apparent diversity of pleasure. The idea here is that there is a *hedonic tone* which accompanies various experiences, and it is in

¹¹³ Sidgwick uses sweetness as an example, while butterflies in the stomach is due to Gosling.

¹¹⁴ See chapter 4 of Bentham's *An Introduction to the Principles of Morals and Legislation*.

virtue of the presence of this tone that we say that an experience is pleasurable. According to Broad,

...there is a quality, which we cannot define but are perfectly well acquainted with, which may be called hedonic tone. It has the two determinate forms of Pleasantness and Unpleasantness... A Pleasure then is simply any mental event which has the pleasant form of hedonic tone, and a pain is simply any kind of mental event which has the unpleasant form of hedonic tone. There is not a special *kind* of mental events, called pleasures and pains; and to think that there is is as if one should solemnly divide human beings into men, women, and blondes. It is of course true that the commonest, and some of the most intense, pleasures and pains are feelings, in my sense of the word. But remorse, which is memory of certain events, having a certain emotional tone, is plainly a pain as much as toothache. And hope, which is expectation of certain events, having a certain emotional tone, is plainly as much a pleasure as the sensation of smell which we get from a rose or a violet (229-30).

On Broad's view, the apparent diversity of various pleasures is due to phenomenological differences other than hedonic tone. For example, headaches and toothaches both share the hedonic tone of unpleasantness, but each has its own specific sensible quality of headachiness and toothachiness, beside further modifications, such as stabbingness, throbbingness, etc., which may be common to both (230). Although toothaches and headaches might seem to be significantly different sorts of mental states, insofar as they are painful, they share the hedonic tone of unpleasantness, and any differences in these experiences are due to other differences in how these experiences feel. Presumably, on Broad's view the same can be said about other sorts of pains, such as sorrow at the death of a loved one or anxiety regarding the prospect of losing one's job. Although there are obvious differences in these various experiences, Broad thinks that they all share a certain hedonic tone, and it is in virtue of this that they are all painful.

Although the hedonic-tone view has the advantage of explicitly acknowledging the apparent diversity of pleasures and pains, while nevertheless holding that there is some unitary experiential property which ties all pleasures (or pains) together, the view is ultimately indefensible. Part of the problem with the account is the obscurity of the notion of hedonic tone. One might take hedonic tone to be a sort of added ingredient in any given experiential state. The idea here would be that when one goes to the dentist one experiences the various sensations associated with such visits, such as the sting of the novocaine needle, the sound of the drill, the soreness in one's jaw from holding it open for a lengthy time, and so forth. The pain consists in the hedonic tone of unpleasantness which is somehow added to these experiences. However, such an account seems false to our actual experience of pain. The sting of the needle is painful in itself, as is the soreness in one's jaw. It simply does not seem to be the case that there is some hedonic tone which accompanies such experiences. Moreover, if there were such a hedonic tone, one would expect it to be easily identifiable via introspection. However, while we might agree that the range of pleasures and pains identified by Broad are genuine pleasures and pains, it is not at all clear that there is some phenomenological feature such as hedonic tone which all pleasures or all pains have in common. In the end, the hedonic-tone theory seems no more able than the Distinctive Feeling view to account for the diversity of pleasures and pains.

These problems with the Distinctive Feeling and hedonic-tone accounts lead fairly naturally toward a very different account of the nature of pleasure, according to which pleasures are those experiential states that we have some sort of pro- attitude toward, and pains are those states we have a negative attitude toward. Call such a view the *attitudinal account of pleasure*. On this view, what all pleasures have in common is a certain relation to

our desires. Broad considers a version of the attitudinal account according to which pleasures are those experiential states which we like and pains are those experiential states which we dislike (237-8). Similarly, Parfit, after dismissing what I have called the Distinctive Feeling view¹¹⁵ on the (by now familiar) grounds of the diversity of pleasurable and painful experiences, he adopts a version of the attitudinal account:

What pains and pleasures have in common are their relations to our desires. On the use of pain which has rational and moral significance, all pains are when experienced unwanted, and a pain is worse or greater the more it is unwanted. Similarly, all pleasures are when experienced wanted, and they are greater or better the more they are wanted. These are the claims of *Preference-Hedonism*. On this view, one of two experiences is more pleasant if it is preferred (493).

Richard Brandt also adopts a version of the attitudinal account, which he calls the *motivational theory of pleasure* according to which an experience is pleasant if and only if it makes its continuation more wanted (40-41).¹¹⁶

Although the allure of attitudinal accounts is understandable given the problems with the Distinctive Feeling and hedonic-tone views, they are nevertheless extremely problematic. The central problem with such accounts has to do with how we are to understand the relevant pro-attitude, whether it be liking, preferring, or wanting an experience. While Broad is quite right to say that finding a certain activity pleasant involves liking the experience, it is less than clear what liking an experience amounts to. One natural reading would be to say that liking an experience or activity entails that experience or activity being accompanied by pleasant feelings or experiences. If we explain Mary's solicitousness in taking care of the

¹¹⁵ Parfit labels it narrow hedonism.

¹¹⁶ Brandt is also skeptical of the Distinctive Feeling (he calls it the *quality of experience view*) because of the apparent diversity of pleasant and painful experiences (35-36).

baby Jesus in terms of her liking the experience, as opposed to doing it out of respect for or fear of God, then we mean that she finds the experience emotionally fulfilling, rewarding, interesting, or in some other way pleasant. However, if we understand liking in this way, liking can hardly be used to explain pleasure, since the point of the analysis was to explain pleasure in terms of liking, and we are now explaining liking in terms of pleasure or pleasantness.

Perhaps Broad's use of the language of liking and disliking was simply unfortunate. We might instead follow Parfit and Brandt in saying that finding experience X pleasant is simply to want or prefer the experience, or perhaps to want it to continue, where this does not assume that the wanting or preferring of X itself involves further pleasant feelings. On this view, to say that one wants, prefers, or likes X is simply to say that one's will is oriented toward X. However, this account of pleasure is open to some serious objections. The most serious problem facing such an account is its vulnerability to a Euthyphro-style objection. This account of pleasure suggests that experiences which are extremely painful would be pleasurable if only our wills were oriented toward them. This makes our judgments about pleasure and pain seem much more arbitrary and optional than they in fact are. Our typical experience of pain and pleasure is that our wills are oriented away from such things as having our fingertips smashed and toward such things as orgasms because such experiences are painful or pleasurable; they are not painful or pleasurable *because* our wills are oriented toward them. Another way of putting this problem is to say that the attitudinal account mistakenly ignores the conceptual possibility that one's will could be oriented toward pain and away from pleasure.

Another problem with defining pleasure in terms of pro-attitudes is that such an

account will mis-classify many painful states as pleasurable ones, particularly if the pro-attitudes are to be given pleasure-free or pleasantness-free analyses. Let's say that a pleasure is any mental or experiential state that the agent has a pro-attitude toward, where this excludes taking-pleasure-in as a relevant pro-attitude for our purposes. There are many painful experiential states which we typically have pro-attitudes toward. If I commit an immoral act such as betrayal of a friend I might very well be overwhelmed by feelings of guilt and remorse. However, that is not to say that I do not approve of my feelings. Indeed, it is quite easy to imagine valuing one's feelings of guilt or remorse and to want them to continue, at least to the degree that the feelings are warranted. The attitudinal account seems to classify such feelings as pleasures, which would be obviously wrong. Similar things can be said about feelings of sorrow at the loss of a loved one, anxiety at the thought of some harm coming to one's child, or anger due to one's unjust treatment. One might approve of such feelings and want them to continue, but that does not mean that such feelings are pleasurable or pleasant.

At this point the supporter of the attitudinal account might try to salvage his view by saying that for an experiential state to be pleasurable it is not sufficient for the agent to want the experience. The agent's desire for the experience must track, or be directed toward, the phenomenological features of the experiential state. Thus, states such as guilt, remorse, sorrow, anger, and anxiety, though desired by the agent, do not count as pleasurable because they are not desired for their phenomenological qualities; they are instead desired insofar as they constitute emotionally appropriate responses to things and events in the world. In other words, such states are desired or wanted *qua* proper emotional response, but not desired *qua* painful experience. Thinking that a state like sorrow or guilt is neither desirable nor desired for its phenomenological qualities is perfectly compatible with thinking that it is both desired

and desirable for its emotional appropriateness, and further that it is desirable for the agent all things considered, in spite of its being a painful experience.

It is possible that such a qualified version of the attitudinal account is what Parfit has in mind when he says that pleasure is *when experienced* wanted and pain is *when experienced* unwanted, the idea being that although I can have pro-attitudes toward painful mental states, these attitudes cannot occur when I am having the painful experience. They can only be something one can have when one reflects on the experience and its relation to things and events in the world. If the idea is that the painful experiences have to go away before one can acquire pro-attitudes toward them, such that one can desire sorrow or anxiety only during those cool moments when one is not afflicted with such mental states and one then reflects on the relation such states have to things, people, and events in the world, then Parfit's claim is simply false. The mother who grieves for her dead child will likely recoil in horror at the prospect of taking a pill or undergoing some therapy which promises to immediately free her of her feelings of grief; thus, it seems perfectly possible to desire to have a painful experience and to want it to continue *when it is experienced*.

There is a second way of reading Parfit's account of pleasure which is more plausible. We might take the claim to be that a pleasure is an experiential state which is desired purely for its phenomenological features, or, we might say, for how it feels. On this view, a desire to retain one's feelings of grief, shame, remorse, or anger does not make such a state a pleasure since the desire is not directed toward the qualitative or phenomenological features of the mental state. Instead, the desire tracks the experiential state's status as an emotionally appropriate response to the world, or, we might say, its status as a virtue or as a virtuous response. Brandt's account of pleasure can also be read in this way:

There is a final feature that should be included in an account of pleasantness. When an experience is pleasant, the (increased) occurrent valence of the continuation of that experience is *causally dependent* on the experiences already going on, say, the taste of ice-cream. We might say that the taste of ice-cream energizes an increased valence of continuing to taste ice-cream, or, at the neo-natal level, energizes the actual retention and swallowing tendencies. To be less metaphorical, we could say that the occurring experience is the *differential cause*, of the increased positive valence of the continuation of that experience (40).

It is important to note the emphasis placed in Brandt's discussion on the role of the *experience* of eating ice-cream in bringing about a desire to continue eating ice-cream.

Contrast the person who eats ice-cream because she enjoys it or finds it pleasurable with the person who eats it out of a sense of moral obligation or perhaps to avoid being rude. Both have a desire to eat ice-cream, but only the person whose desire to continue eating ice-cream tracks the phenomenological features of her experience of eating ice-cream can be said to *enjoy* the experience or find it *pleasurable*.

Although this modified version of the attitudinal account, according to which a pleasure is any experiential state desired by the agent for how it feels or for its phenomenological qualities, solves the problem of the mis-classification of states such as sorrow, grief, guilt, and remorse as pleasures, it is still open to the Euthyphro-type worries I raised earlier. For it seems conceptually possible that a creature's will could be systematically oriented toward those mental states which we take to be painful and away from those mental states we take to be pleasurable, and for such a creature's desires to track the phenomenological qualities of those states—that is, for it to desire or to be averse to those states for the way they feel. The most natural way of describing such a creature is as desiring that which is intrinsically undesirable, and being averse to that which is intrinsically

desirable, or as desiring pain and being averse to pleasure. However, the attitudinal account suggests that the painful states are, for the creature in question, pleasures, which is implausible. One might think that there is something deeply confused about the notion of a creature which systematically desires pain and is averse to pleasure. Indeed, it is difficult to *imagine* such a creature, but this is due mainly to ubiquity of pleasure-seeking and pain-avoiding behavior. All of the people and (sentient) animals we come into contact with exhibit such behavior, and given our own experience with pleasurable and painful mental states it is hard to make sense of a creature whose desires run in the opposite direction. Moreover, given the correlation between painful mental states and serious structural damage to the body, it is quite doubtful that any such creature could in fact survive for long. However, that is not to say that we cannot conceive of such a creature. We can, and the attitudinal account of pleasure and pain says that the states such a creature desires are in fact pleasures, which is clearly the wrong thing to say about such states. One is averse to having one's hand placed on a hot stove because the burning sensation is painful; the sensation is not painful *because* one has an aversion to it.

Although there might be other possible accounts of pleasure, these views I have discussed seem to be the main contenders for a theory of pleasure. The Distinctive Feeling view, notwithstanding its virtues of simplicity and unifying power, seems obviously false. There simply is no one mental state or event which can be called pleasure. The hedonic-tone view ultimately fails in its attempt to combine a recognition of the diversity of pleasures and pains with the simple, unifying property of positive or negative hedonic tone; to the extent that the view is distinguishable from the Distinctive Feeling view it is obscure, and to the extent that it isn't distinguishable it simply seems false. The attitudinal account of

pleasure is either circular or implausible. It is circular if the relevant pro- and negative attitudes are defined in terms of pleasure and pain. If they are understood in terms of the bare orientation of the will, then the account has some extremely counterintuitive implications, such as the mis-classification of painful experiential states such as grief, sorrow, anxiety, and anger as pleasures so long as they are desired or approved of. This particular problem can be avoided if we define pleasure as an experiential state which is desired by the agent for the way it feels. Although this qualified version of the attitudinal account is more plausible, it is still open to Euthyphro-type objections.

Thus, we are led to the conclusion that the right view to have about pleasure and pain is that some states are intrinsically desirable while others are intrinsically undesirable, with the former being pleasures (broadly construed) and the latter pains. This way we avoid the notion that something can be made pleasurable by the orientation of one's will, while avoiding the indefensible commitment to something like hedonic tone. This is not to say, of course, that the same type of activity will be equally pleasurable to different people. I enjoy eating hotdogs and you don't; you enjoy riding the Ferris Wheel but I don't; Larry enjoys listening to hip-hop music but Steve doesn't. There is obviously much diversity in people's likes and dislikes, and it might be thought that my intrinsic-desirability account of pleasure will have a difficult time accounting for this. However, we must keep in mind the various things that can contribute to someone's liking or disliking something. For example, perhaps the reason you don't like hotdogs has something to do with certain associations you have with them, such as that they remind you of a hot-dog eating contest you once witnessed which you found especially disgusting, while part of the reason I like them is my association of them with childhood picnics or trips to the baseball game. Another possibility is that we simply

have certain physiological differences such that hotdogs taste different to us, much in the same way that orange juice tastes different depending on whether one has just brushed one's teeth. Perhaps the reason I don't like the Ferris Wheel is that I panic in high places; this makes riding a Ferris Wheel a very different experience for me. Similarly, perhaps the reason Larry likes hip-hop and Steve doesn't is that Larry is more familiar with the origins of many of the samples in hip-hop songs, and thus the music has an emotional resonance for Larry it wouldn't have if he shared Steve's musical background. Thus, the notion that certain experiential states are either intrinsically desirable or undesirable is perfectly compatible with the recognition of the rather common phenomenon of different people enjoying different things or the same person finding a given activity pleasant at one time but painful or unpleasant at another.

III: Is Pleasure Merely a Conditional Good?

In this section, I want to consider the claim that pleasure is a merely conditional good. It is undeniable that pleasure sometimes can be a bad thing. Obviously, it can be instrumentally bad—e.g., the pleasure one experiences upon trying heroine can lead to a heroin addiction, which can lead to all sorts of bad things both for the addict and for those who come into contact with him. It also seems that there is something bad about pleasure which is taken in the suffering or misfortune of others. Moreover, there seems to be something bad about undeserved pleasure. How do these considerations bear on our conviction that pleasure is intrinsically good? Are they genuine counterexamples to the claim? The fact that pleasures can be instrumentally bad (or that instances of pain can be instrumentally good) is not at all surprising, and is no threat to the intrinsic goodness or

desirability of pleasure; indeed, the phenomenon of something being at once both intrinsically good and instrumentally bad is fairly common, and any plausible account ought to recognize this. The other cases, however, are more problematic. In this section, I shall consider the following theses regarding the alleged conditionality of the goodness or desirability of pleasure:

1. Pleasure or enjoyment is intrinsically good, valuable, or desirable only if it is pleasure or enjoyment taken in some sort of excellent thing or activity.
2. Pleasure or enjoyment is intrinsically good, valuable, or desirable only if it is morally innocent.
3. Pleasure or enjoyment is intrinsically good, valuable, or desirable only if it is morally deserved.

Each of these has a certain *prima facie* plausibility. Also, each seems to call into question the claim that pleasure is intrinsically good. However, all of these claims are false, at least when the issue is whether pleasure is intrinsically good *for the agent*. A consideration of the issue in light of our earlier discussion of the nature of pleasure will help to show why even pleasures taken in non-excellent activities, undeserved pleasures, and sadistic pleasures are beneficial to the agent, at least when considered solely *as pleasures*.

Take first the proposition that only pleasures taken in excellent things or activities are intrinsically good or desirable. Robert Adams argues for a view along these lines, claiming that well-being consists in the enjoyment of the excellent (93-101). On his view, enjoyment or pleasure is a necessary but not sufficient condition for some thing or activity to contribute to one's good. Derek Parfit suggests a similar view near the end of *Reasons and Persons*:

We might claim that what is best for people is a composite. It is not just their being in the conscious states that they want to be in. Nor is it just their having knowledge, engaging in rational activity, being aware of true beauty, and the like. What is good for someone is neither just what Hedonists claim, nor just what is claimed by Objective List Theorists. We might believe that if we had *either* of these, *without the other*, what we had would have little or no value. We might claim, for example, that what is good or bad for someone is to have knowledge, to be engaged in rational activity, to experience mutual love, and to be aware of beauty, while strongly wanting just these things. On this view, each side in the disagreement saw only half the truth. Each put forth something as sufficient that was only necessary. Pleasure with many other kinds of object has no value. And, if they are entirely devoid of pleasure, there is no value in knowledge, rational activity, love, or the awareness of beauty (502).

Both Adams and Parfit think that subjective goods such as pleasure and objective goods such as knowledge and rational activity are merely conditional goods. Pleasure without excellence is worthless, as is excellence without pleasure or enjoyment. This is a position which has a certain prima facie plausibility. A life completely devoid of pleasure, even it contains a good number of perfectionist goods, sounds like a rather bleak existence. Similarly, in evaluating a life, we are not just concerned with how much pleasure is experienced; it is also important to know what sorts of things the agent is taking pleasure in and the degree to which she exemplifies perfection or excellence in what she does. However, one can accept these fairly modest claims without accepting the stronger claims of Parfit and Adams that pleasure or enjoyment without excellence is *completely worthless*, and that excellence or perfection without pleasure or enjoyment *has no value whatsoever*. These stronger claims are implausible, and ought to be rejected.

Certain pleasures, such as that derived from listening to fine music, working a difficult math problem, successfully making a challenging mountain climb, and caring for

one's child are indeed cases of enjoyment of the excellent, and would count as contributing to one's good or well-being on the Adams/Parfit view. However, there are numerous other pleasures which are not taken in excellent things or activities which nevertheless contribute to the well-being of the agent. For example, lazily dozing in a hammock on a warm afternoon and feeling the cool ocean breezes is clearly something which can make one's life better. However, it would stretch the meaning of the terms "perfection" and "excellence" beyond recognition to view this sort of activity as an exemplification of excellence or perfection, or as some sort of significant achievement. Similarly, the pleasure one gets from eating an ordinary but nevertheless tasty piece of cake is the sort of thing that can make one's life better, but presumably we would not want to identify the piece of cake as an excellent thing for the purposes of an Adams/Parfit analysis of well-being; for if the cake is admitted as an excellent or objective good, then the list of excellent things will be so lengthy that there will be very few pleasures which will fail to contribute to one's good. However, denying the objective goodness or excellence of the cake would lead Parfit or Adams to the implausible conclusion that the enjoyment or pleasure one gets from the cake in no way contributes to one's good or well-being.

At this point, the defender of the conditionality of the goodness or desirability of pleasure might respond by admitting that the pleasure derived from a piece of cake or from an afternoon in the hammock contributes to one's good, but nevertheless insist that for pleasure to be intrinsically good or desirable it must be *morally innocent*. In defending such a view, one might appeal to our intuitions about sadistic pleasures. Again, there is something *prima facie* plausible about the claim that sadistic pleasures are not intrinsically valuable or desirable. However, we need to be careful about what conclusions we draw about the way in

which such states are undesirable or disvaluable. Certainly, considered *as pleasures* sadistic pleasures are indeed valuable and intrinsically desirable to the agent. This follows from our definition of pleasure as any experiential state which is intrinsically desirable for its phenomenological content, or for how it feels. Thus, even pleasure taken in the suffering of others is intrinsically valuable, so long as the relevant phenomenological content is present; for it is the phenomenological content which makes the state of pleasure intrinsically desirable *qua pleasure*.

Similar conclusions will follow if we accept one of the other conceptions of the nature of pleasure discussed earlier. If pleasure is taken to be some sort of distinctive feeling or hedonic tone, and this distinctive feeling is caused by some sadistic activity or by viewing the suffering of others, or the hedonic tone is present in the experiential states caused by such things, then even sadistic pleasures will have to be admitted as being intrinsically desirable for the agent. For on these conceptions of pleasure, it is the distinctive feeling or hedonic tone which is intrinsically desirable. If the distinctive feeling or hedonic tone is present, then we have an instance of pleasure, which is intrinsically desirable regardless of its origins or object. The attitudinal account of pleasure, which seems to be the most serious rival to my account, also suggests that pleasure is not merely conditionally good, and that even sadistic pleasures are intrinsically good *qua pleasures*. For the attitudinal account identifies pleasure as any experiential state which is intrinsically desired by the agent for its phenomenological content, or for how it feels. On this view, it is the agent's attitudinal states—i.e., his pro- and con attitudes—which determine whether a particular mental state is a pleasure or a pain. However, for such states to be plausibly identified as pleasures or pains, the relevant pro- and con attitudes must track the experiential content of the mental states, and not their status as

appropriate emotional responses to the world. In other words, the relevant attitudes must track the *feeling* of the mental state, and not its status as an instance of a virtuous or vicious response to the world, if it is to constitute a pleasure or pain. If we accept the attitudinal account's contention that having a pro- attitude toward a given mental state (where this attitude tracks the experiential content of the state) is enough to render that mental state an instance of pleasure, and further that such states contribute to the good of the agent in virtue of such an attitude, then it follows that even sadistic pleasures contribute to the good of the agent.

A consideration of the status of pain as an unconditional *bad* for the agent can shed further light on the status of pleasure as an unconditional good. On my conception of the nature of pleasure and pain, certain experiential states are intrinsically undesirable due to the way they feel, and these are pains. On the attitudinal account, there are certain experiential states which agents have negative attitudes toward due to the phenomenological content of these states, and the presence of such content, along with the relevant negative attitude, is enough to make such states intrinsically bad for the agent. However, on either view, if there is something to be said in favor of the painful mental state – e.g., that it is an emotionally appropriate response to the world or that it is the result of a heroic act of self-sacrifice – it does not follow that the painful state is not a burden or a cost to the agent, when considered *as a pain*. Indeed, the reason we think of certain actions as cases of self-sacrifice is our commitment to the notion that certain states, regardless of their moral worthiness, are, *qua experiential states*, intrinsically undesirable for the agent, and detrimental to his or her well-being. Since the positive moral status of an instance of significant pain is not sufficient to render it beneficial to the agent *qua painful experience*, by parity of reasoning, it follows that

the negative moral status of a pleasure is not sufficient to render it of no value to the agent *qua pleasurable experience*.

The notion that sadistic pleasures necessarily contribute to the well-being of the sadist is admittedly troubling. However, it is important to note the reasons we can offer to a sadist for changing his behavior. First, we can emphasize his moral (as opposed to prudential) reasons for changing his behavior, particularly if his disposition to find pleasure in the suffering of others leads him to engage in harmful acts toward others; there is no reason to think that all reasons for action must be reasons of self-interest. Second, we can emphasize to the sadist the fact that there are intrinsic goods other than pleasure, such as the virtue of interpersonal sympathy. The fact that we would not want sadism to be exemplified by someone we care about, such as a child or close friend, is evidence that we recognize sadism as something which detracts from the well-being of the sadist, at least insofar as it constitutes vicious moral character. Thus, although the sadist gets something good when he tortures someone (viz., pleasure), he pays a price in exemplifying a vicious moral disposition. If we are to think that it is never the case that sadistic pleasure benefits the agent all things considered, this can only be because we think that the price he pays in moral virtue is so high that it cannot be outweighed by the pleasure he gets from the act of torture. While such a claim about the relative value of pleasure and the virtue of impersonal sympathy might be true, it is no threat to the claim that pleasure (*qua pleasure*) is an intrinsic, unconditional good for the agent.

A defender of the conditionality of the value of pleasure might appeal to the notion that pleasure or happiness is bad if it is undeserved. Kant famously relies on such an intuition in his defense of the claim that the only unconditional good is the morally good will:

It is impossible to think of anything at all in the world, or indeed even beyond it, that could be considered good without limitation except a good will. Understanding, wit, judgment and the like, whatever such *talents* of mind may be called, or courage, resolution, and perseverance in one's plans, as qualities of *temperament*, are undoubtedly good and desirable for many purposes, but they can also be extremely evil and harmful if the will which is to make use of these gifts of nature, and whose distinctive constitution is therefore called *character*, is not good. It is the same with *gifts of fortune*. Power, riches, honor, even health and that complete well-being and satisfaction with one's condition called *happiness*, produce boldness and thereby often arrogance as well unless a good will is present which corrects the influence of these on the mind and, in so doing, also corrects the whole principle of action and brings it in conformity with universal ends – not to mention that an impartial rational spectator can take no delight in seeing the uninterrupted prosperity of a being graced with no feature of a pure and good will, so that a good will seems to constitute the indispensable condition even of worthiness to be happy (4:393).¹¹⁷

Much of what Kant says here deals with the ways in which certain character traits and attributes such as courage, health, and understanding can be used for evil ends, and there is nothing especially controversial or surprising in the observation that traits which are normally thought of as virtuous or beneficial can sometimes be instrumentally bad. What is controversial and interesting for our purposes is Kant's claim that an impartial rational spectator can take no delight in seeing happiness attach to a morally bad will, and that a morally good will is the indispensable condition...of worthiness to be happy.

Ross makes a similar argument about the relation between pleasure and moral virtue in support of his claim that the allocation of pleasure to the virtuous is an intrinsic good and deserves to join virtue, pleasure, and knowledge on his list of fundamental goods:

¹¹⁷ In citing Kant's *Groundwork* I make use of the standard Academy citation system found in most translations.

In so far as the goodness or badness of a particular pleasure depends on its being the realization of a virtuous or vicious disposition, this has been allowed for by our recognition of virtue as a thing good in itself. But the mere recognition of virtue as a thing good in itself, and of pleasure as a thing *prima facie* good in itself, does not do justice to the conception of merit. If we compare two imaginary states of the universe, alike in the total amounts of virtue and vice and of pleasure and pain present in the two, but in one of which the virtuous were all happy and the vicious miserable, while in the other the virtuous were miserable and the vicious happy, very few people would hesitate to say that the first was a much better state of the universe than the second. It would seem then that, besides virtue and pleasure, we must recognize...as a third independent good, the apportionment of pleasure and pain to the virtuous and the vicious respectively (1930: 138).

Ross and Kant share the intuition that it is a bad thing for the morally vicious to be happy, or to have pleasure. Also, both are offering a judgment about the goodness of a certain *state of affairs* namely, that of the morally vicious possessing happiness or pleasure. I think that Kant and Ross are right that the possession of pleasure and happiness by the morally vicious, and pain and suffering by the morally virtuous, is a bad state of affairs, and that there is a sense in which morally undeserved happiness or pleasure is bad. However, we cannot conclude that the happiness or pleasure is bad *for the agent*. Indeed, the judgment that such a state of affairs is bad depends on the fact that happiness and pleasure benefit the agent. If they did not, then it would make no sense to speak of their possession by the morally wicked as undeserved. If the pleasure found in immoral actions were of no real benefit to the agent, then we could say that such an agent gets what he deserves i.e., he would get something of no value to him. Moreover, unless we assume that pleasure and happiness are intrinsically valuable to the agent *qua experiential states*, then we can make no sense of the claim that it is a bad state of affairs for pleasure and happiness to be allotted to the morally vicious. Thus,

the sorts of considerations offered by Kant and Ross in no way undermine the claim that pleasure is an intrinsic good for the agent, and actually offer support for the proposition that it contributes to the well-being of the agent.

While it is true that there is something bad about a morally wicked person gaining pleasure or happiness from his wicked actions, it would be intellectually dishonest to insist that it is bad in the sense of being bad for the wicked agent. Indeed, the intuition that it is especially bad for a person to profit from his morally vicious actions presupposes that what he gains is in fact profitable to him. In this section I have argued that pleasure and freedom from pain are, when considered solely as experiential states, beneficial to the agent. However, I have indicated that this is only one aspect of the good, and that there are intrinsic goods other than the possession of, and freedom from, certain kinds of experiential states. One of these goods is an important aspect of moral virtue what I refer to as having the appropriate emotional response to things in the world. In the next section, I argue that such a disposition is an important aspect of the good, and thus that there are things other than pleasure and freedom from pain which can make someone's life go well.

IV: Having the Appropriate Emotional Response

In attacking the notion that desirable mental states such as pleasure, enjoyment, and happiness are the only thing which can make a life go well, I will make the case for the intrinsic value of states such as virtue (understood in a certain way), knowledge, understanding, and rationality. I will start with virtue and discuss the others in the next section. The kind of virtue that I have in mind involves being disposed to respond to various things and events in the world appropriately. This means taking pleasure in the right things,

and being pained by the right things. My sense is that philosophers, partly due to their natural commitment to the pursuit of knowledge and understanding, are already inclined to be skeptical of hedonistic-type views. Thus, they tend naturally to gravitate toward positions which are I think unduly dismissive of the intrinsic value of pleasure and related states, no matter how implausible such views are. Although there may be marginally more support for something like hedonism or a quality-of-experience view about the good among non-philosophers, I think that even here there is precious little support for such a view. Virtually everyone is committed to the proposition that there is more to life than pleasure and enjoyment, and I want to make use of some sources from popular culture to help illustrate and flesh out the pluralist account of the good which I want to defend.

In an episode of *The Sopranos* called 'The Happy Wanderer,' mob boss Tony Soprano complains to his psychiatrist Dr. Melfi about his sometimes overwhelming sense of anger. Who, she asks, is he angry at?

I don't know who the fuck I'm angry at. I'm just angry, okay? Why the fuck am I here? I even asked to come back. I got the world by the balls and I can't stop feeling like I'm a fucking loser.

When Melfi asks if it's his mother that makes him feel like a loser, Tony insists that it isn't:

No, we've wasted enough oxygen on that one. It's everything and everybody. I see some guy walking down the street, you know, with a clear head. You know the type. He's always fucking whistling like the happy fucking wanderer. I just want to go up to him and I just want to rip his throat open. I just want to fucking grab him and pummel him right there for no reason. Why should I give a shit if a guy's got a clear head? I should say, 'Ah, salut; good for you.'

Tony is clearly puzzled by his own feelings toward the 'happy wanderer.' On the one hand, Tony seems to resent him, and wants to rip his throat open. Part of his resentment is no

doubt due to the fact that the happy wanderer makes Tony feel inadequate or a failure in some way. Yet, Tony also realizes that there is something deeply unreasonable about his resentment toward the happy wanderer. He describes himself as wanting to pummel him for no reason, and wonders why he should give a shit if a guy's got a clear head. Tony seems to realize that happiness is not a zero sum game, and that the clearheadedness of the happy wanderer is not, in and of itself, something that should make Tony angry or resentful. Tony realizes that, if anything, the appropriate response is to be happy for the happy wanderer and to wish him well.

Although one might take the exchange between Tony and Melfi simply as evidence of Tony's inherent viciousness, I think that it is fairly easy to sympathize with Tony in his feelings toward the happy wanderer. For we might interpret Tony's complaint as an unusually violent articulation of the mixed feelings we often have toward those who are too easily pleased or who are happy even though there are good reasons to be unhappy. On the one hand, because we think that happiness is intrinsically desirable, we think that the happy wanderer's life is going well insofar as he is happy. Tony experiences this as feelings of inadequacy i.e., as feeling like a loser because he realizes that the happy wanderer has something valuable, and this valuable thing is something that Tony lacks. On the other hand, Tony feels a certain contempt for the happy wanderer, for his happiness seems to be grounded in a certain obliviousness or indifference to the bad or evil things in the world. There is something foolish, silly, or ridiculous about the happy wanderer, as is illustrated by the lyrics of the *Happy Wanderer* song which runs through Tony's mind at the end of the episode:

I love to go awandering
 Along the mountain track
 And when I go, I love to sing
 My Knapsack on my back

Valderee, valderah,
 Valderee, valderah-ha-ha-ha-ha-ha
 Valderee, valderah
 My knapsack on my back

I love to wander by the stream
 That dances in the sun
 So joyously, it calls to me
 Come join my happy song

Valderee, valderah,
 Valderee, valderah-ha-ha-ha-ha-ha
 Valderee, valderah,
 Come join my happy song

Although the happy wanderer enjoys a carefree, pleasant existence, there is something frivolous and childish about his approach to life. He is not a serious person, and to the extent that his happiness is bought at the price of failing to confront the difficulties and everyday ugliness of our world, he is both foolish and irresponsible in his reactive attitudes, and thus is an object of Tony's contempt or indignation. In short, Tony responds with a mixture of envy and contempt for the happy wanderer: envy of the happiness that he enjoys, but contempt for his failure to respond emotionally in a way that is appropriate to a serious adult.¹¹⁸

Tony Soprano is not the only one to feel something like contempt for someone like the happy wanderer. In an episode of *Curb Your Enthusiasm*, Larry David expresses similar feelings about Ted Danson. After Larry and his wife Cheryl spend an enjoyable evening bowling with Ted and his wife Mary they both agree that Mary is great, but Larry

¹¹⁸ It is significant that the Happy Wanderer song goes through Tony's head just at the point when he has done something which forces his daughter to face certain facts about her father's lifestyle. It comes out more explicitly in a later episode that Tony was at least partially motivated by the desire that his daughter not continue to be under certain illusions about the way in which he makes his living. It is also significant that he continues to feel quite conflicted about his actions in this regard, expressing regret that this required rubbing his daughter's nose in the ugly truth.

announces that Ted is someone he could do without. When Cheryl asks what's wrong with Ted, Larry says that he's a little strange. Specifically,

Everything's heaven with him. The gum. He had a piece
of gum. Oh this is heaven. Had a taste of a chocolate bar
Oh! Oh! I'm in heaven. The *parking space* is heaven.
It's all heaven.

Cheryl responds to this diatribe by saying, "Well, he's a happy guy; he *enjoys* things." Larry seems unsure how to respond to this, other than to shrug as if to say, "Yeah, I guess." Again, I think that Larry's reaction to Ted is something one can sympathize with. On the one hand, he realizes that Cheryl has a point that Ted is a happy guy, and that his happiness is partly grounded in a capacity to take pleasure in everyday things; this is good for Ted so far as it goes. On the other hand, there is something disconcerting about someone who is genuinely so promiscuous or indiscriminating in his pleasures. Although Ted's happiness, *qua* happiness, makes his life go better, the fact that his responses are so out of proportion to the objective value of the things in which he finds pleasure seems to detract from Ted's life. In other words, there is more to leading a good life than having positive feelings such as pleasure or happiness; it is also important that our emotional responses are appropriate in the sense of corresponding to the real value of the things in which we take pleasure or find happiness.

In J.D. Salinger's *Catcher in the Rye*, Holden Caulfield expresses a similar sentiment regarding some acquaintances of his:

Brossard and Ackley both had seen the picture that was playing, so all we did, we just had a couple of hamburgers and played the pinball machine for a little while, then took the bus back to Pencey. I didn't care about not seeing the movie, anyway. It was supposed to be a comedy, with Cary Grant in it, and all that crap. Besides, I'd been to the movies with Brossard and Ackley before. They both laughed like hyenas at stuff that wasn't even funny. I didn't even enjoy sitting next to them at the movies (48).

Like Tony Soprano, Holden Caulfield is a somewhat troubled individual, but I think we can sympathize with his attitude toward people like Brossard and Ackley. Holden is bothered by people laughing at things that aren't funny, and this could be because he perceives it as another instance of the phoniness about which he continually complains. While it is true that people will sometimes emit phony laughter as a way of ingratiating themselves with others (particularly with those in positions of authority), this is not at all typical of the sort of laughter one finds in movie theaters. Such laughter is typically genuine or authentic. However, we might still be bothered by even genuine laughter if we think that there is nothing objectively funny about what's being laughed at. To the extent that Holden has a legitimate complaint against or criticism of Brossard and Ackley, it is due to the fact that they seem to be responding inappropriately to what is on the screen, and I think that anyone who has found himself in a room full of people who are laughing (genuinely) at something that simply seems not to be funny can sympathize with Holden's annoyance with his two acquaintances. Nevertheless, there is a sense in which we can be envious of people like Brossard and Ackley because they have the capacity to laugh at and enjoy things that we cannot, and to that extent, their lives go better than ours. However, that does not mean that we want to be like Brossard and Ackley, *all things considered*. For the additional pleasure or enjoyment that they derive from bad movies (or whatever) has to be weighed against the intrinsic badness of their inappropriate responses. An indiscriminating palate is not something that we want for ourselves, nor is it something that we want for those we care about, even if such a palate does indeed result in greater pleasure or enjoyment.

This distaste for the idea of enjoying or taking pleasure in things that are not worthwhile or objectively valuable is not confined to the faculty club or graduate student

lounge; it also extends to the frat house. There are certain virtues to which the frat guy aspires, and certain ideals to which he is committed. One virtue is coolness, and the corresponding ideal is what one might call The Cool-Guy Aesthetic. But what precisely is a Cool Guy? One important feature of the Cool Guy is that he tends to lack enthusiasm. There is much that he is blasé about, and he is not easily excited. Boys often start to become aware of the demands of The Cool-Guy Aesthetic around the onset of adolescence; this typically involves the loss of many childhood enthusiasms and an increasing tendency to affect a kind of boredom with, or disinterest in, many of the things one formerly found enjoyable. While some of the newfound lack of interest is no doubt genuine, it is also the case that much of it seems simply to be an affectation—a response that is grounded in an increasing awareness that some things are not worthy of one's attention, even though they may be quite enjoyable. Thus, a young teenager, though he may still have the capacity to have a good deal of fun at Disneyland or watching cartoons, might resist the temptation if he thinks of such enjoyment as kids' stuff.

By early adulthood, The Cool-Guy Aesthetic typically has taken hold, and it is further developed and reinforced in the culture of the frat house. One might think that the prevailing culture of the frat house would be hedonistic, but that would be an oversimplification, at best. In the frat house, pleasure and enjoyment are important, but one must be careful not to find pleasure or enjoyment in the wrong things. Sex is obviously thought to be something worth taking pleasure in, but it has to be the right sort of sex. Gay sex, for instance, is definitely not cool, no matter how much one might enjoy it. Also, one must be wary of falling deeply in love, even with a member of the opposite sex; for the intense and overpowering feelings associated with being in love are an almost certain recipe for losing one's cool. But one must

be especially careful not to take excessive pleasure in academic or other intellectual pursuits.

Here's how a character in Tom Wolfe's *I Am Charlotte Simmons* describes the demands of

coolness:

The cool guy doesn't flatter anybody or act obsequious or even impressed by somebody unless it's some athlete...and you don't act enthusiastic unless it's about sports, sex, or getting high. It's okay to be enthusiastic about something like Dickens...or Foucault or Derrida for that matter but if you want to be cool, you don't show it, you don't say it, you don't even let on. A cool guy and I've seen this happen can secretly work his ass off five no, four nights a week at the library, but he has to make light of it if anybody catches on. You know what the favorite major of the cool guy is? Econ. Econ is fireproof, if you know what I mean. It's practical. You can't possibly be taking it because you really love economics (409-10).

Although Wolfe's book is technically a work of fiction, it is meant to be a more-or-less realistic description of life on the modern campus. The author spent a good deal of time researching his book by attending various fraternity and sorority parties, and I have to say that his description of the frat-house milieu fits my own (admittedly) limited experience. What is most interesting about his description is that, if there is a commitment to hedonism among the frat guys, it is a hedonism which is qualified in some way. Clearly, something's being pleasurable or enjoyable is not taken to be sufficient for its being good or desirable, all things considered.

It is not at all clear how the hedonism if it can be called that of the frat guys is to be understood, or how it is to be qualified. There are a couple of things a philosophically-inclined frat guy might say. First, he might say that pleasure and enjoyment are good, but they are merely *conditionally* good, the idea being that such states are good for the agent only if they are found in objectively valuable activities, such as getting high, having heterosexual

intercourse, watching the Big Game, etc. As I have argued, it seems extremely implausible to claim that the goodness of pleasure or enjoyment is conditional in this way; surely, if I derive enjoyment from one of the forbidden activities such as studying Econ, then that pleasure benefits me. A second, and I think more plausible, response the frat philosopher might give is to say that, while pleasure and enjoyment are indeed intrinsically valuable, these are not the *only* good things. It is also important that we have the right sorts of dispositions, and that we desire the things which truly are worth desiring. In short, pleasure and enjoyment are not all that one needs for a good life. One must also (says the frat philosopher) be cool, where this means having the right sorts of dispositions, desires, and emotional responses to things in the world. Unfortunately, these goods can conflict, so that the cool life is not necessarily the most pleasant, joyful, or the happiest life.¹¹⁹

The philosophy professor and the frat guy will often disagree about what counts as something worth pursuing or taking pleasure in. Similarly, we may or may not agree with Holden Caulfield about what is truly funny and thus worth laughing at. We will also likely disagree at times with Tony Soprano regarding the conditions under which it is appropriate to whistle a happy tune, and perhaps with Larry David regarding the degree of pleasure or enjoyment appropriate to finding a choice parking space. But while we may disagree on the details, it is important to recognize a certain convergence of opinion. On the one hand, it seems undeniable that happiness, pleasure, enjoyment, and related states are intrinsically valuable. On the other hand, the notion that such good feelings are the only thing to make a life go well is a view that very few people actually hold. Virtually all of us are committed to

¹¹⁹ This tension between coolness and happiness is further illustrated by the fact that people who appear on the cover of *Rolling Stone* magazine virtually never appear happy.

the proposition that it is a bad thing to find too much enjoyment in things or activities which are not really worthwhile. Not only do we want ourselves and those we care about to have a good time and enjoy things, we also want our reactions to be appropriate. We want to have discriminating palates. We do not want to play the rube or the happy wanderer even if such a role is quite pleasant or enjoyable.

The capacity to respond appropriately to things and events in the world is not just a matter of experiencing pleasure and related states at the right times and in the right ways; it also extends to negative states such as pain, sadness, anxiety, fear, anger, and so forth. Although these states are all intrinsically undesirable *qua* experiences, it does not follow that we should want to completely rid ourselves of them, nor is it the case that we would want those we care about to be completely free of such experiences. Of course, some of this can be attributed to the *instrumental* value of such states; for example, fear or anxiety might play a role in helping one to avoid certain dangerous situations. However, there also seems to be something *intrinsically* desirable about having the capacity to experience certain states. For instance, if a loved one is in danger, then a certain level of anxiety seems to be the appropriate response, even if there is every reason to believe that such anxiety is of no instrumental value.

Take the case of Bob and Mary, who are having their first child. Bob is present for the delivery, and is there mainly for emotional support. As is often the case, there are various complications in the delivery. The labor is long, and there is talk of resorting to a c-section. The baby's heart rate is alarmingly low at times. The nurse tells Mary to be sure to take deep breaths with her contractions so that the baby gets plenty of oxygen. The nurse also says something about the umbilical cord being wrapped around the baby's neck. On top of all of this, Mary is clearly in a good deal of pain and discomfort. Bob's natural response to this is

to be worried and anxious for both mother and child. His anxiety has little or no instrumental value, since there is precious little that Bob can *do* about the situation; if anything, his anxiety might be instrumentally disvaluable, since it may very well cause Mary to be anxious.

Moreover, Bob's anxiety is not intrinsically valuable, *qua* experience; he would not want to feel such anxiety for the sake of being anxious. Nevertheless, there is something valuable about his negative feelings, and I take it that most of us, if we were in Bob's place, would want to react in the way that he does. Moreover, I take it that we would want those we care about to have Bob-like reactions to similar situations, even though the reactions consist largely in certain bad — i.e., unpleasant — feelings. Thus, Bob wants little Bob Jr. to grow up to be the sort of person who experiences feelings of pain and anxiety when those close to him are in pain, danger, or distress. Importantly, this desire for Bob Jr. to have a capacity for such states is a manifestation of Bob's caring for Bob Jr., which suggests that this capacity is partly constitutive of the good of Bob Jr.

Next, take the example of Fred and Ethel, who were married for over 50 years. When Lucy goes to Ethel's funeral, she is somewhat shocked to see Fred seated conspicuously next to the coffin, weeping loudly and copiously. Part of what shocks Lucy is what she takes to be the unseemliness of Fred's behavior; it's disconcerting to see Fred in such a state, with such a loss of self-control. But part of her surprise is also due to the fact that this behavior seems rather out of character for Fred, for Fred and Ethel seemed to spend most of their time insulting and needling one another. Lucy had no idea that Fred loved Ethel this much. Indeed, this spectacle causes Lucy to have deeply ambivalent feelings toward Fred. On the one hand, she has lost respect for him due to his loss of self-control. But she also pities Fred, because she knows the burden of the pain that he feels. However, she also has a newfound

respect and affection for Fred due to the deep love he had for Ethel, and she thinks that it is, in some sense, a good thing that Fred responds to Ethel's death in the way that he does.

Driving home from the funeral, Lucy tries to make sense of her conflicted feelings regarding Fred. How does she reconcile her disappointment with Fred's loss of self-control with her respect for his pained response to Ethel's death? One possibility would be to identify self-possession, understood as a certain emotional imperviousness to the difficulties and vicissitudes of life, as a genuine intrinsic good which can sometimes conflict with the capacity to be deeply moved by the death of a loved-one. Such a view might be not only acceptable but quite attractive within a pluralistic framework. However, even if one is resistant to the idea of self-possession, so construed, as having intrinsic value, this case offers a clear example of the conflict between the value of having the capacity to feel pain at the loss of a loved-one and the value (or disvalue) of certain mental states. Lucy recognizes that the pain that Fred feels at the passing of his wife is a bad thing for Fred to have to endure. Certainly, Lucy would not want Fred to have to experience it for the sake of the experience. However, she also thinks that Fred is in some sense better off for having such a capacity. Fred's life would be worse if he were the sort of person who lacked such an emotional response; that is, the capacity to feel pain in certain circumstances is partly constitutive of Fred's good, even though the exercise of this capacity results in feelings which, considered simply *as feelings*, are bad for Fred. Thus, Lucy concludes that it is a good thing, all things considered, for Fred to have such dispositions, and she vows to instill such dispositions in her own son, Little Ricky, even though this will no doubt result in some emotionally difficult times for Little Ricky down the road; for Lucy realizes that there is more to living a good life than pleasure and the avoidance of pain.

Anger offers another illustration of this sort of conflict of values. Anger is a disutility; it is unpleasant and not worth pursuing for itself *qua* experience. If you set up shop in the night club and try to sell an anger pill, you will have a very difficult time competing with the Ecstasy dealer. Of course, some might be interested in your anger pill for its instrumental value; for instance, it might be helpful if one plans to get into a fist fight. Nevertheless, I take it that, like the capacity to experience emotional pain and anxiety, the capacity to feel anger is something which we do value intrinsically, at least in certain circumstances. The reason we value our anger in certain situations is that our anger seems to be the appropriate response. Take Joe for instance. Joe gets angry at some situation and has an outburst. Sally tells Joe that he shouldn't be angry, because it doesn't do any good to get angry. Sally continues, your getting angry doesn't change anything, other than making things unpleasant for yourself and those around you. Joe responds quite reasonably, we might think that the point of getting angry is not always to change the situation (though it can have this effect). Joe does indeed value his anger, but he values it because he takes anger to be the appropriate emotional response to an infuriating situation. This is not to say that he doesn't recognize that his anger is a disutility, and thus is bad for him in some sense, nor does he deny that it can be instrumentally bad. However, Joe does not want to be the sort of person who does not get angry when anger is the most appropriate response. He agrees with Aristotle that people who are not angered by the right things, or in the right way, or at the right times, or toward the right people, all seem to be foolish (NE 1126a).

I've been emphasizing some of the ways in which one might refuse to sacrifice the good of having the appropriate emotional response for pleasure or enjoyment. However, one should not conclude that it is *never* reasonable to sacrifice one's capacity to respond

appropriately for something like happiness or freedom from pain. These are distinct values, and they sometimes conflict, and though having the right sort of emotional capacities is very important, it does not have anything like infinite value or lexical priority. If there is enough at stake in terms of pleasure, happiness, contentment, pain, anxiety, and so forth, it might very well make the most sense to sacrifice the capacity to respond to things in the world with the degree of pleasure or pain appropriate to their objective or genuine value. For instance, although we do indeed value our tendency to respond to infuriating situations with anger, if one is constantly surrounded by genuinely infuriating situations, things, and people, so that one is constantly experiencing feelings of warranted anger, then it might make the most sense to work to weaken this capacity for anger. The anger itself is unpleasant and thus intrinsically undesirable, so that a large enough quantity of it could very well outweigh the value of the capacity.

The capacity to feel pain or distress at the thought of the pain of others or at the loss of a loved one is another valuable capacity which we might very well be better off without under certain circumstances. If there seems no end in sight to the husband's intense grief or distress at the loss of his wife, then we might think he would be better off, all things considered, without such a capacity. We might even suggest that he take some sort of numbing medicine to reduce his feelings of grief. Similarly, although we do value the capacity or disposition to feel pain or distress at the pain of others—indeed, sympathy is one of the most important of the virtues—we do not cling to it come what may. If I am surrounded by a tremendous amount of pain or suffering (say I am sent to a concentration camp) then it might very well be better for me to rid myself of the capacity for sympathy. For the pain and distress I feel through such a capacity might be so overwhelming that it

outweighs the benefit I get from having the appropriate emotional response to all of the suffering around me. However, this is not to say that the husband who takes the numbing medicine and the prisoner in the concentration camp do not lose something of genuine value, for one cannot help but regret the loss of such dispositions. We typically think of such dispositions as an important aspect of our humanity, and it is natural to see the loss of them with its attendant desensitization to the suffering of others as a form of dehumanization.

Similarly, just because one becomes convinced that the things one enjoys are not really worthy of the attention of a serious, sophisticated adult it does not necessarily follow that one should work to rid oneself of the capacity to enjoy such things. If the pleasure or enjoyment one finds in bad television or music is intense enough, and one finds it difficult enough to develop other, more sophisticated tastes, then it might make the most sense to continue enjoying watching *Gilligan's Island* and dancing to *The Macarena*, no matter how foolish or ridiculous it renders one. Of course, the development of one's palate typically involves not only the loss of the ability to enjoy certain things, but also the gaining of new tastes, so that one can now enjoy reading Proust and dancing in the ballet. However, there is no guarantee that this will be the case, and it would not be that surprising to find that the person of poor taste is nevertheless much happier and leads a much more pleasant life than the aficionado of French literature and Russian ballet. And it would not be that surprising to find that some who attempt to develop their palates end up in an unfortunate middle point, where they have largely rid themselves of the capacity to enjoy *Gilligan's Island* and *The Macarena*, but have not developed a love of Proust and ballet. In such cases, one must weigh the good of having good (or improved) taste against the good of the pleasure one gets from various sources, and it could turn out that, for some people, they are better off, all things

considered, if they stick to television and mediocre pop songs.

V: Knowledge, Understanding, and Rationality

Many philosophers have claimed that knowledge, understanding, and rationality are intrinsically good. As we saw in previous chapters, such states typically play a central role in perfectionist accounts of the good, sometimes threatening to eclipse other important goods. My view is that knowledge is intrinsically valuable, but that it sometimes conflicts with other goods such as pleasure and proper emotional responsiveness. In this section I make the case for the intrinsic value of knowledge, understanding, and rationality and discuss some of the ways in which these various goods can conflict. In arguing for the intrinsic value of states such as knowledge and rationality, I make constructive use of the views of Aristotle along with more recent perfectionists such as T.H. Green, and Thomas Hurka . Although the accounts put forth by Green and Hurka are ultimately untenable as comprehensive accounts of the human good, they nevertheless offer useful analyses of practical rationality, along with some compelling reasons for thinking that the exercise of our capacities for practical and theoretical rationality must be an important part of our good.

Why think that knowledge is intrinsically good? Although the intrinsic value of states such as knowledge and understanding might ultimately be undeniable, this is not *obviously* the case. There are a number of things, for instance, that a hedonist might say about the instrumental value of knowledge, and a sophisticated hedonist might try to account for our intuitions about the apparent intrinsic value of knowledge and understanding by pointing to the feelings of satisfaction or pleasure which typically follow the satisfaction of our desires, including our desire for knowledge. She will also point to the various ways in

which knowledge and understanding can direct us to greater sources of pleasure and help us to avoid sources of pain and suffering. Since knowledge and understanding are often directly and instrumentally pleasurable, and ignorance often leads to pain, one might be tempted to conclude that our intuitions about the intrinsic value of knowledge and understanding are mistaken: it might be thought that if a case of knowledge were not pleasurable, and if it truly had no power to lead us to pleasure or the avoidance of pain, then it would be of no value at all.

Although it is true that knowledge and understanding are often instrumentally valuable, and ignorance can often lead to pain and suffering, we should not exaggerate these tendencies. First of all, the feelings of satisfaction which typically accompany our desire for knowledge require not knowledge for their fruition so much as a mere feeling of conviction, or the belief that a certain important truth has been discovered. Second, we should not lose sight of the fact that knowledge is sometimes quite painful, and ignorance is sometimes bliss. Although the blissful wise man is a bit of a stock character from the intellectual traditions of both the East and West, the tortured existentialist thinker who faces (what he takes to be) the cold, hard fact of a Godless, meaningless universe is an equally recognizable image, as is the scientist who knows about the coming of some great catastrophe but lacks the power to do anything to prevent it. Moreover, experience teaches us that the ignorant do not always pay for their lack of knowledge or understanding with pain and suffering; such people can live quite pleasurable and enjoyable lives. While the hedonist might be willing to bite the bullet and say that a feeling of conviction is every bit as good as genuine knowledge or that the blissfully ignorant are clearly better off than those who know the unpleasant truth, such a response is ultimately unsatisfactory. The instrumental value of knowledge in terms of

pleasure and pain cannot adequately explain or justify our intuitions about the value of knowledge and understanding. It seems that we are committed to the proposition that they are intrinsically valuable.

Ross offers a (by now familiar) thought experiment to convince us that knowledge deserves a place on his list of intrinsic goods, alongside pleasure, virtue, and the apportionment of pleasure to the virtuous:

It seems clear that knowledge, and to a less degree what we may for the present call right opinion, are states of mind good in themselves. Here too, we may, if we please, help ourselves to realize the fact by supposing two states of the universe equal in respect of virtue and of pleasure and of the allocation of pleasure to the virtuous, but such that the persons in the one had a far greater understanding of the nature and laws of the universe than those in the other. Can anyone doubt that the first would be a better state of the universe (1930: 138-139)?

Ross thinks that the intrinsic value of knowledge is self-evident, which is not to say that he thinks its intrinsic value is obvious. His thought experiment is meant to bring to our attention the intrinsic goodness of knowledge. If we imagine two possible states of the universe which are identical in every way, except that one contains much more knowledge than the other, and ask ourselves which is better, Ross thinks that we must agree that the one with greater knowledge is a better state of the universe. The hedonist, by contrast, has to say that the two are equally good, since they are identical in terms of the amount and distribution of pleasure and pain.

I agree with Ross that, of the two states of the universe, the one with greater knowledge is better than the other. I also agree that this suggests that we are committed to thinking that knowledge is intrinsically valuable. However, it is not clear from Ross's thought experiment that we have to conclude that knowledge is *good for* those who possess it.

All Ross's thought experiment commits us to is the proposition that the universe with greater knowledge is a better state of affairs than the other. This leaves open the question whether the state of affairs is better in virtue of the fact that knowledge contributes to the good or well-being of those who possess it. We might adopt the sort of view held by Thomas Nagel, who thinks that certain goods, which he calls perfectionist, have an objective, intrinsic value which is independent of their value to the individuals who experience or use them:

By [perfectionist ends or values] I mean the intrinsic value of certain achievements or creations, apart from their value to individuals who experience or use them. Examples are provided by the intrinsic value of scientific discovery, of artistic creation, of space exploration perhaps. These pursuits do of course serve the interests of the individuals directly involved in them, and of certain spectators. But typically the pursuit of such ends is not justified solely in terms of those interests. They are thought to have an intrinsic value, so that it is important to achieve fundamental advances, for example, in mathematics or astronomy even if very few people come to understand them and they have no practical effects. The mere existence of such understanding, somewhere in the species, is regarded by many as worth substantial sacrifices. Naturally opinions differ as to what has this kind of worth. Not everyone will agree that reaching the moon or Mars has the intrinsic value necessary to justify its current cost, or that the performance of obscure or difficult orchestral works has any value apart from its worth to individuals who enjoy them. But many things people do cannot be justified or understood without taking into account such perfectionist values (129-30).

Nagel goes on to say that certain difficult or challenging projects, such as climbing Mt. Everest or mastering the *Well-Tempered Clavier*, can take on enormous significance for us once we adopt them as our own. Such projects, once adopted, make autonomous claims on us...which they need not have made in advance (130). Nagel thinks that it makes perfect sense for the man who has dedicated himself to mastering the *Well-Tempered Clavier* to refuse to go out to dinner on the grounds that he must stay in to practice, but it would be

strange for him to say that he had to master the *Well-Tempered Clavier* (130).

I take Nagel to be claiming that the existence or occurrence of certain achievements such as someone's mastering the *Well-Tempered Clavier*, climbing Mt. Everest, or making an important scientific discovery is an objectively good thing. It is objectively a better state of affairs if such achievements take place. However, for the climbing of Everest or the mastering of the *Well-Tempered Clavier* to contribute to any particular person's good — i.e., for such activities to be *good for* someone — these perfectionist goods must play some sort of role in that person's goals or projects. Nagel suggests that the mere existence of certain kinds of knowledge or understanding, somewhere in the species, is worth significant sacrifice, but he seems to think that such sacrifice is to be justified by appeal to the intrinsic value of the knowledge, and not to any benefit such knowledge confers onto the possessor of the knowledge. This is an important claim for our purposes, since what we are after is an account of the personal good. We want to know whether a perfectionist good such as knowledge can be objectively good, in the sense of contributing to the overall value of the universe, without contributing to the good of any particular person. Does it make sense to think that a perfectionist good such as knowledge can be worth pursuing or promoting purely due to its intrinsic value, and completely independently of any value it might have for persons?

There are certain goods which seem to fit Nagel's notion of a perfectionist value. Great works of art and certain objects of natural beauty seem to have a value or worth that is independent of any value they might have for persons. Thus, I think that Moore is right that, if we are to compare the value of two possible worlds, both of which are never to be enjoyed or experienced by any intelligent agent, where one world is exceedingly beautiful, including whatever on this earth you most admire — mountains, rivers, the sea; trees, and sunsets, stars

and moon...all combined in the most exquisite proportions, so that no one thing jars against another, but each contributes to increase the beauty of the whole, while the other is nothing but a heap of filth, containing everything that is most disgusting to us...without one redeeming feature we are not indifferent between these two.¹²⁰ It makes a difference which world obtains, even in the absence of intelligent beings, which suggests that certain objects have a value which is independent of any value they might have for persons. This assumes, of course, an account of beauty according to which it is constituted by certain properties and relations such as order and proportionality. Although such an account is controversial, it seems plausible enough when applied to certain objects of artistic and natural beauty, such as musical compositions and ecosystems. One might reasonably think that a complex ecosystem such as the Amazon Rain Forest or the Serengeti has value simply in virtue of the way in which its components work together, and that the *Well-Tempered Clavier* is similarly valuable due to its inherent complexity, order, and proportionality. A world with such objects of beauty seems preferable to a world of chaotic filth.

Although the inherent beauty of a complex ecosystem or musical composition offers an explanation of how it is that such things can have objective, perfectionist value in Nagel's sense, it is less clear how knowledge, the acquisition of knowledge, the climbing of Everest, or the mastering of the *Well-Tempered Clavier* can also be said to have such independent value. While the existence of knowledge somewhere in the species might be a good thing, it would seem very odd to explain its value by appeal to its *beauty*; indeed, beauty and knowledge seem quite distinct categories, with each having its own sort of value. Similarly, there does not seem to be anything particularly beautiful about the *acquisition* of knowledge

¹²⁰ Moore 1903a (83-84).

or the climbing of Everest. Although the *Well-Tempered Clavier* might be quite beautiful, it is not at all clear that the *mastering* of the work is itself a thing of beauty. To explain the value of these other perfectionist goods, we have to appeal to the difficulty one encounters in the achievement of such goods. Successfully climbing Everest, mastering the *Well-Tempered Clavier*, and understanding a difficult scientific theory are all challenging activities, each requiring the development and exercise of important human capacities. If such activities are to have perfectionist value, it will be in virtue of their requiring the use of such important human capacities.

Although it might be plausible to say that objects of natural or artistic beauty such as an ecosystem or the occurrence (as opposed to the mastery) of a musical composition are objective goods in the sense that Nagel intends, the analogy between such works of art and objects of natural beauty on the one hand and human activities and achievements on the other is imperfect to say the least. Even if we think that the Serengeti or the occurrence of the *Well-Tempered Clavier* is, because of its inherent beauty, objectively good independently of the ways it is *good for* anyone, we should not conclude that activities such as climbing Everest, mastering the *Well-Tempered Clavier*, or obtaining some bit of important knowledge or understanding can also be good without being *good for* someone. What we need is some explanation for why someone's climbing Everest, mastering the *Well-Tempered Clavier*, and understanding the laws of physics are objectively good things — why it is a better state of affairs for such achievements to exist or take place. Note that we have a working explanation for the objective value of the Serengeti and the occurrence of the *Well-Tempered Clavier* — the inherent beauty of these objects. Fortunately, the perfectionist has something to say on this score. She can say that it is a better state of affairs for such achievements to take place

because those who climb Everest, master the *Well-Tempered Clavier*, or understand the laws of physics benefit from such achievements in significant ways. Such achievements are not just things that happen. They are things done by certain people, and if we think that climbing Everest or understanding physics is a genuinely excellent achievement, then we should conclude that those who achieve these things do in fact benefit from them. They benefit insofar as they exemplify excellence in the execution of such difficult tasks. If climbing Everest or understanding physics is a genuine perfectionist good due to the way it requires the excellent use of one's capacities and abilities, then any achievement of such a perfection will be not only objectively good, but also *good for* the person who achieves the perfectionist good, because it is his perfection which is being exemplified.

A modified version of Ross's thought experiment can give further support to the claim that knowledge is intrinsically valuable, not only objectively, but also *for* the possessor of knowledge. Recall that Ross asks us to imagine two states of the universe which are equal in all respects, except that one contains more knowledge than the other. We agreed with Ross that the state of the universe with greater knowledge is better. However, I argued that the most plausible reason for thinking it better is that knowledge contributes directly to the well-being of those who possess it. Now, instead of imagining two states of the universe, imagine two possible courses of life for someone you care deeply about, such as a child, spouse, a very close friend, or perhaps a future time-slice of yourself. The two courses of life promise identical amounts or degrees of pleasure, freedom from pain, and whatever other items you think are essential components of a good life (other than knowledge). The only difference is that in one course of life your loved-one will achieve significant levels of knowledge and understanding about a wide range of subject matter, while the second course of life will entail

tremendous ignorance and a woeful lack of understanding. Which would you choose or recommend for your loved-one?

I take it that virtually all of us, assuming that we are truly convinced that the two courses of life are equally good in all respects other than the possession of knowledge, would choose the life of knowledge and understanding for our loved-one. The fact that we are not indifferent between these two possible courses of life indicates that we take knowledge to be not merely instrumentally good, but intrinsically good as well. The fact that knowledge is something we would choose for someone we care about is evidence that we take knowledge to be not just good objectively, but *good for* our loved-one. This latter claim is also supported by the fact that we are concerned not just with the total amount of knowledge in possible states of affairs, but also with its distribution. Assume that we can hold the current total amount and distribution of goods other than knowledge and understanding constant, and then improve the state of the universe by adding some amount of knowledge and understanding. It matters to us whether the knowledge and understanding will be possessed by those we care about, those we are indifferent toward, or those whom we positively dislike. The fact that we would prefer the knowledge and understanding to go first to those we care about, those we are indifferent toward second, and those we dislike third (or perhaps not at all) suggests that knowledge and understanding are the sorts of things we take to contribute to the well-being of those who possess them.

Although these considerations reveal our commitment to the proposition that knowledge and understanding are constituents of what we might call the *personal good* that is, we take them to contribute directly to the well-being of those who possess them some difficult questions remain. The first has to do with the place of practical rationality in our

account of the good. So far, we have been concerned to establish the intrinsic value of theoretical rationality, and have all but ignored the status of practical rationality. The second question has to do with our reasons for thinking that theoretical rationality is constitutive of the personal good. The skeptic will be quick to point out that arguments demonstrating our commitment to X are not equivalent to arguments for the truth of X . What we need is some sort of explanation for *why* we are committed to the intrinsic value of theoretical and practical rationality. That is, we need to say something about our normative reasons for thinking that a world or a life containing knowledge, understanding, and rationality is better. These two questions are connected. In order to explain why theoretical and practical rationality are intrinsically good for creatures like us, we have to understand the role of our capacity for agency both in explaining our nature and in explaining and justifying the special dignity that human beings are thought to possess.

In Book X of the *Nicomachean Ethics* Aristotle tells us that the pursuit of theoretical study is what is best and most proper for us. Although we might be skeptical of Aristotle's metaphysical claims about the proper function of human beings, along with his rather strong claims about the supremacy of theoretical contemplation, there does seem to be something right about his linking of rational activity with human nature. While we might not be *essentially* rational, it nevertheless seems an important fact about us that we are typically capable of various kinds of rational activity. Although there is no reason to think that this sort of rational activity is unique to human beings, no other beings that we know of have this ability to the extent that humans do. And, even if it turns out that there are beings which surpass us in the capacity for rationality, this would not seem to alter the fact that our rationality is an important human trait—so important in fact that it is what ultimately grounds

our notion that human beings have a special dignity which is lacked by other members of the animal kingdom, and is what makes us worthy of a special kind of respect.¹²¹

Numerous commentators on Aristotle have pointed out that there seems to be significant tension in Aristotle's thinking on the human good.¹²² In Book X of the NE we seem to get the view that theoretical contemplation is the sole end that human beings should pursue. Theoretical contemplation is the supreme good, and happiness or eudaimonia extends just so far as study extends. Aristotle tells us that the more one studies the happier he is (1178b30-32).¹²³ Moreover, at least some of what Aristotle says in Book X suggests that theoretical contemplation possesses the characteristic traits of true happiness such as self-sufficiency and completeness.¹²⁴ Such comments seem to suggest a rather narrow form of perfectionism, according to which the only intrinsic good is something like knowledge or theoretical understanding. However, in other parts of the NE, especially Book I, Aristotle suggests a broader or more inclusive conception of the human good. While this more inclusive conception of the good also gives pride of place to theoretical contemplation, it seems to leave room for the intrinsic value of other forms of virtuous activity. On this more inclusive understanding, the good consists in rational activity in accord with virtue, where this

¹²¹ This tying of the special status and dignity of human beings to our capacities for rational deliberation is evidenced in the role that such capacities play in current debates about the moral permissibility of abortion and various forms of euthanasia.

¹²² A tremendous amount of literature has been produced over the last few decades regarding the relation between what Aristotle says about the good in Book I and what he says in Book X. Gavin Lawrence's *Aristotle and the Ideal Life* contains a useful survey of the literature. For a classic treatment of the issue see Hardie's *The Final Good in Aristotle's Ethics*.

¹²³ Irwin's translation.

¹²⁴ See especially 1177a-1178b.

involves not just theoretical contemplation but also the proper ordering of our desires and actions in accord with the demands of practical rationality. Integral to the exercise of such rationality is understanding the proper role of virtues such as wisdom, justice, temperance, courage, and generosity and acting in accord with this understanding.

It is this latter, more inclusive Aristotelian conception of the human good which Green builds on in the development of his conception of the good as self-realization. Although I am very skeptical of the notion that the human good can be understood solely in terms of self-realization as Green understands it,¹²⁵ I nevertheless think that Green's position is partially right and offers a compelling explanation for why rationality is something that we ought to value as an intrinsic good. Recall that Green starts with the concept of a person, and identifies personhood with our capacities for both theoretical and practical deliberation. Unlike brutes and children, who are slaves to their doxastic impulses, normal adult humans are able to distance themselves from such impulses and form and act on beliefs which are grounded in and supported by the available empirical evidence and the rules of logic. This ability to distance ourselves from and resist our doxastic impulses is what makes us epistemically responsible agents. It is an essential part of our personhood, and thus the exercise of such personhood requires the development and exercise of our powers of theoretical deliberation.

There is more to our personhood, however, than the exercise of our capacities for theoretical deliberation; practical rational deliberation is also essential. This involves the ability to distance ourselves from our immediate, first-order desires and to form a conception of the good which can guide our actions. Just as the ability to distance ourselves from our

¹²⁵ See chapter 4 of this dissertation.

doxastic impulses is what makes us epistemically responsible agents, it is this ability to distance ourselves from and critically evaluate our desires, thereby forming a conception of the good according to which we can make choices, that makes us practically responsible agents. This capacity for practical rational deliberation is essential to our personhood, and the exercise of our personhood thus requires the development and exercise of such capacities for practical rational deliberation.

It is not hard to see what the development and exercise of our capacities for theoretical rational deliberation should consist in. At its most refined levels it will consist in the sorts of epistemic practices associated with the various sciences. However, it will also involve more pedestrian factual beliefs about such things as the location of the post office and the winner of last year's Super Bowl. Hurka offers a sophisticated account of the relation between isolated, particular bits of knowledge and the more general claims associated with, for instance, scientific theories.¹²⁶ On Hurka's view, not all pieces of knowledge are equally valuable. A fundamental law of the universe is more valuable than the knowledge of the winner of last year's Super Bowl because of the former item's greater generality. The fundamental law of the universe explains a wide range of phenomena and thereby exhibits what Hurka calls *extent*. Since many other bits of knowledge (or beliefs) can be derived from this fundamental law of the universe, it can also be said to exhibit a great deal of *hierarchical dominance*. Theoretical rationality is largely a matter of developing systems of individual beliefs or pieces of knowledge where these systems are arranged in complex structures which exemplify explanatory extent and dominance. Only human beings who have reached a fairly high level of intellectual maturity are capable of devising such complex doxastic systems, and

¹²⁶ See Hurka 1993 (114-120).

this ability seems to be one of the things which ground our special status and dignity.

Hurka's account of theoretical rationality, though somewhat sketchy as I present it here, gives us a tolerably clear understanding both of the nature of theoretical perfection and why it is valuable. However, as we saw in our previous discussion of Green's view, it is a bit harder to specify what counts as proper practical rational deliberation. Nevertheless, I think there are some fairly uncontroversial things we can say about practical deliberation. The most obvious thing is that such deliberation requires giving thought to one's options in life, familiarizing oneself with one's alternatives, critically reflecting on the origins of one's desires along with how these desires are related to one another, thinking about alternative ways of ordering one's desires, investigating how other people form conceptions of their own good and the extent to which their choices result in the attainment of presumed goods such as pleasure or achievement, and so forth. In short, practical rational deliberation requires a significant amount of effort so that one can, in some sense, transcend the immediate desires with which one is faced at a particular time. This is largely what it means to be a person, and to the extent that we are concerned with our personhood we should be concerned with the development and exercise of our capacities for rational deliberation—both theoretical and practical.

This final point, of course, raises the question of exactly how concerned we should be with our personhood. As I argued in a previous chapter, I don't think there is any reason to think that any of us is essentially a person. Nevertheless, it seems to be an important fact about human beings that most of us are indeed persons, possessing those capacities for rational deliberation which are constitutive of personhood. Thus, *qua person*, one always has reason to further develop one's capacity for rational deliberation. However, we are not

merely persons; we also have a certain amount in common with the non-persons of the world. Like non-rational animals, we are capable of pleasure and pain, and while we are presumably subject to a much wider range of emotions, feelings, and sensations than most other animals, we nevertheless share certain basic utilities and disutilities, such as the pleasures of eating and sex along with physical pain and certain basic emotions such as fear. Thus, we might say that *qua animal* or *qua sentient being* I always have reason to seek pleasure and avoid pain. Since the development of my personhood sometimes requires the sacrifice of pleasure or can even be painful, the goods associated with these two aspects of my being or nature sometimes come into conflict.

Some might argue that this way of understanding practical rationality, where the development and exercise of its constitutive capacities is thought of as one of many competing goods, is seriously confused. One might insist that it is the resolution of just such conflicts that practical rationality is charged with. On this view, the exercise of our practical rationality involves weighing the value of goods such as pleasure and contentment against the value of knowledge, understanding, and the sorts of practices I have associated with the exercise of our capacity for rational deliberation. I think that there is something right about this conception of practical rationality, but that it ignores an important distinction involving the ways in which we use and conceive of practical reason. We might label these *formal* and *substantive* aspects of practical reason. By formal aspects I have in mind just those practices I identified as constituting the exercise of practical rationality: thinking about options, investigating alternative conceptions of the good, reflecting on one's desires, etc. The substantive aspect of practical reason involves weighing the value of the exercise of such capacities against the value of other goods, such as pleasure, emotional fulfillment, proper

emotional responsiveness, and so forth. One might achieve a great deal along the formal dimension of practical rationality but score fairly low along the substantive dimension. This would be the case if someone spent *too much* time thinking about what she ought to do all things considered, such that she actually achieved very little in the way of other, competing goods. Even though her life is not going that well, all things considered, we should not say that there is *no* value in the time she spends contemplating the good and the right. For such reflection is an important part of being a person, and is one of the things that sets us apart from the lower animals. Qua person, our obsessive contemplator surely achieves *something* of intrinsic value, though her life is not as good as it might be all things considered.

What we would like, of course, is a theory about the substantive aspect of practical reason — one which would give us some definitive answers about how these various goods such as pleasure, proper emotional responsiveness, theoretical rationality, and formal practical rationality might be balanced against one another. Unfortunately, I think the prospects of such a theory are pretty bleak. While we certainly do make judgments about such goods in cases where they conflict, it seems extremely difficult, if not impossible, to give any precise rules about how the various goods are to be weighed against each other. Similarly, the idea of attaching some sort of quantitative measure to the various goods seems to border on the absurd. The best we can do, I think, is demonstrate that there are various goods such as pleasure and freedom from pain, proper emotional responsiveness, theoretical rationality, and formal practical rationality. We can also see that these various goods are grounded in the complex nature of human beings — specifically, in the sentient, emotional, conative, and rational aspects of our nature. Thus, we have a fairly rich mix of goods which are given some unity by the idea that the good of a being is ultimately grounded in the sort of

being that it is. Moreover, it is a conception of the good which takes a realistic view of human nature, admitting that we have much in common with the lower animals, but at the same time giving our higher faculties their just due. While such a view compromises a certain amount in the realm of determinacy and theoretical simplicity, it nevertheless offers a reasonably unified and coherent account of the human good.

VI: Are Perfectionist Goods Conditional on Endorsement?

Earlier in this chapter, I considered whether pleasure is a merely conditional good, and argued that it is not. The same question can be raised about perfectionist goods such as the exercise of our capacities for practical and theoretical deliberation, proper emotional responsiveness, the creation of great works of art, athletic accomplishments, and so forth. Whereas the claim about the conditionality of pleasure is usually that it has to have an objectively valuable source or at least that it must be morally innocent, the claim about perfectionist goods is that they cannot contribute to one's personal good or well-being unless one desires them, takes pleasure in them, or endorses them. In this section I will argue that there is no reason to think that perfectionist goods are conditional in this way. I begin with a discussion of the view that the good consists in enjoyment of the excellent or in finding pleasure in objectively good things. I then move onto a critical discussion of what is often called the endorsement constraint, according to which some activity or way of life can contribute to the good of the agent only if he endorses it.

Recall the view suggested by Parfit that what's good is taking pleasure in things that are objectively valuable; on this view, taking pleasure in an activity is a necessary condition for its contributing to one's personal good or well-being. Adams' conception of the good as

the enjoyment of the excellent shares this commitment to the conditionality of perfectionist goods. How plausible are the views of Parfit and Adams? Parfit's view has a certain prima facie plausibility due to the important role played by pleasure in any defensible account of the good. Also, the fact that people do tend to find pleasure in the successful pursuit of objectively valuable ends lends a certain degree of credibility to the view. Moreover, the image of a person finding pleasure, contentment, and tranquility in the successful pursuit of objectively valuable activities such as raising a family, practicing a worthwhile craft, achieving some significant theoretical understanding, or making a positive and lasting contribution to one's community seems to fit very closely our conception of an ideal life, which further supports Parfit's account of the good.

However, just because an ideal life consists in the combination of certain goods such as pleasure and the pursuit or exemplification of excellence, it does not follow that each component or constituent of the good is not intrinsically valuable in isolation from the others. Indeed, if we think that a life of pleasure found in objectively good activities is better for the agent than a life of the same degree of pleasure found in activities lacking in objective goodness, there must be something about the objectively good activities—some underlying good-making properties—which somehow contribute to the well-being of the agent. If these underlying properties can contribute to the well-being of the agent when pleasure is present alongside them, it is very difficult to see why they should fail to contribute to the well-being of the agent when the pleasure is absent. For example, let's assume that what makes a life of pleasure found in the mastering and performing of the works of the great composers such as Bach, Mozart, and Rachmaninoff better for the agent than a life of pleasure found in the mastering and performing of the works of Brittany Spears, Ozzy Osbourne, and Vanilla Ice is

the greater complexity and inherent beauty in the works of the great composers: the works themselves are more beautiful and their mastery requires the development of one's musical capacities at a higher level of complexity and sophistication. It's possible that someone could develop these capacities to a very high level without deriving much pleasure from them, and in such a case the good-making properties of complexity and beauty are present, though pleasure is lacking. In such a case, we should say that the pleasureless exemplifying of the capacities associated with the mastering of the great composers, along with the production of the beauty inherent in the works, makes some significant contribution to the well-being of the agent. This is further supported by a consideration of a case where an agent simply lacks a capacity for significant amounts of pleasure or joy. In this case, the agent will not experience any real pleasure from the mastering and performance of either the great composers or the mediocre pop icons. However, I think we would say that a joyless mastering of the great composers contributes more to one's life than a joyless mastering of bad pop music; we are not indifferent between the two options. This suggests that objectively valuable activities can contribute to one's well-being even if they are not sources of pleasure.

The Parfit/Adams account of the good also has a difficult time making sense of proper emotional responsiveness as a constituent of the good, particularly when the proper response is a disutility such as pain, grief, sorrow, or anger. Let's assume that experiencing such emotions is a type of excellence on the Parfit-Adams view. In order for such responses to contribute to the good of the agent who experiences them, she would have to enjoy them or find pleasure in them. However, it's hard to see how this can be the case in normal circumstances. If my child dies then, if I am emotionally healthy, I will experience sorrow and grief. I have argued that the dispositions associated with such a response are something

which we value, and the fact that we want those we care about, including ourselves, to have such dispositions and responses is evidence that we take such responses to be partly constitutive of the personal good. But on the Parfit/Adams view, responding with grief and sorrow contributes to the agent's well-being only if she enjoys or takes pleasure in such feelings. Clearly, these are not the sorts of things we can or should enjoy or take pleasure in. Thus, the personal good cannot be adequately explained in terms of enjoyment of the excellent or taking pleasure in objectively valuable things.

The defender of the conditionality of perfectionist goods might, at this stage in the argument, admit that it makes little sense to speak of the emotionally healthy person *enjoying* or taking *pleasure* in his feelings of grief, sorrow, or anger but still insist that for these experiential states, or any other perfectionist goods for that matter, to contribute to one's good or well-being one must have some sort of pro-attitude toward them. The relevant pro-attitude might be taken to be one of *endorsement*. While it is hard to make sense of a parent enjoying or taking pleasure in his feelings of grief over his dead child, we can easily make sense of him endorsing those feelings and the dispositions and capacities underlying them. We can also easily imagine the attitude of endorsement extending to other experiential states such as anger, fear, contentment, and many others. Endorsement can also extend to other goods such as knowledge, athletic achievement, the practice of a craft, and so forth. The idea would be that, while it is not necessary for knowledge to contribute to one's well-being that one enjoy it or take pleasure in it, one must nevertheless endorse it if it is to contribute to one's good. Since one is *ceteris paribus* likely to endorse what one enjoys or finds pleasure in, there will often be a connection between the two; however, one is not forced to endorse one's enjoyments or pleasurable responses or to reject one's feelings of pain, sorrow, or

anger.

Call the view that endorsement is a necessary condition for something to contribute to one's personal good the *endorsement constraint*. Should we accept it? Although the endorsement constraint has some genuine advantages over the Parfit/Adams view, it is nevertheless ultimately indefensible. The first thing to note about the endorsement constraint is that it specifies a necessary and not sufficient condition for something to be a constituent of someone's personal good. This is as it should be since the mere fact of endorsement seems clearly insufficient to allow something to make someone's life better. However, by making endorsement a necessary condition we are left with an extremely demanding criterion which will have a difficult time accounting for many of our considered judgments about the ways in which a person's life can be made better. Since the endorsement constraint is held by some prominent philosophers such as Will Kymlicka and Ronald Dworkin, my strategy in the rest of this section will be to consider the arguments they put forth in support of endorsement as a necessary condition for something's contributing to one's personal good.

Kymlicka argues that a life cannot be made better by being led from the outside according to values the agent doesn't endorse. As Kymlicka puts it,

...while we may be mistaken in our beliefs about value, it doesn't follow that someone else, who has reason to believe a mistake has been made, can come along and improve my life by leading it for me, in accordance with the correct account of value. On the contrary, no life goes better by being led from the outside according to values the person doesn't endorse. My life only goes better if I'm leading it from the inside, according to my beliefs about value. Praying to God may be a valuable activity, but you have to believe that it's a worthwhile thing to do—that it has some worthwhile point and purpose. You can coerce someone into going to church and making the right physical movements, but you won't make someone's life better that way. It won't work, even if the coerced person is mistaken in her belief that praying to

God is a waste of time. It won't work because a life only goes better if led from the inside (and some values can only be pursued from the inside) (12).

Kymlicka argues here that paternalism is self-defeating. The only way to improve someone's life is by getting them to change their behavior in light of their own value commitments. However, this rather sweeping claim seems to have some counterexamples. One can certainly make the life of a child better by coercing him to behave in certain ways, regardless of whether the child's behavior is being led from the inside. Moreover, the example Kymlicka uses here is quite limited in application. He uses it to support his claim that paternalism is self-defeating, and this certainly seems to be the case when it comes to prayer. After all, those who would coerce us into prayer presumably are concerned not just with whether we make the external movements associated with prayer, but also that we have those experiential states which the devout typically enjoy, such as reverence toward God. There is no reason to think that coerced prayer will have such a result; indeed, outright coercion (as distinguished from more subtle forms of manipulation) are probably counterproductive in getting people to adopt the feelings of reverence which the supporter of prayer is after. However, that is not to say that all perfectionist goods share this feature with religious conviction. A paternalist might be satisfied with getting people to do certain kinds of excellent things, regardless of whether these people are leading their lives from the inside. If I think that knowledge of physics or mastery of the *Well-Tempered Clavier* makes someone's life better, it is not at all clear why my attempts at coercing people to acquire the relevant factual knowledge or know-how must be *self-defeating*. Although such coerced excellence may ultimately be an indefensible violation of people's autonomy, such attempts at getting people to exemplify excellence do not have to be self-defeating.

Dworkin shares Kymlicka's views about the self-defeatingness of paternalism, but employs a good deal more analytical machinery in defending this view. Since my purpose in discussing Dworkin is mainly critical, we need to be clear about the details of his view. So I will quote extensively from Dworkin's discussion of the endorsement constraint in *Sovereign Virtue*. Dworkin makes three interrelated distinctions in arguing for the endorsement constraint: (1) volitional vs. critical interests; (2) the challenge model vs. the impact model; and (3) additive vs. constitutive accounts of the relation among an agent's desires, activities, and well-being. According to Dworkin, in order to answer the question, "What makes a life a good or successful one?" we must eschew what he calls the reductionist impulse of utilitarians who see the good solely in terms of desire satisfaction, and recognize the need for complexity and structure within the idea of well-being:

We must recognize, first, a distinction between what I shall call volitional well-being and critical well-being. Someone's volitional well-being is improved, and just for that reason, when he has or achieves what in fact he wants. His critical well-being is improved by his having or achieving what it makes his life a better life to have or achieve. Sailing well and freedom from dentistry are part of my own volitional well-being: I want them both, and my life therefore goes better, in the volitional sense, when I have them. I take a different view of other things I want: having a close relationship with my children, for example, securing some success in my work, and what I despair of obtaining some minimal grasp of the state of advanced science of my era. These I regard as critical interests because I believe that my life would be a less successful one if I failed to have, or wholly failed to achieve, these goals (242).

This distinction between critical and volitional interests is important for Dworkin because it captures our sense that one's desires are typically a reliable guide to facts about one's good or well-being while at the same time recognizing that an agent can be mistaken about her good or well-being. As Dworkin puts it: "People can fail to recognize their own critical interests.

It makes sense to say that someone who has no regard for friendship or religion or challenging work, for example, leads a poorer life for that reason, whether he agrees or not. We also make critical judgments about ourselves; people all too often come to think, toward the end, that they have ignored what they only then realize is really important to their lives (216).

Dworkin's second distinction describes two different ways of understanding our critical interests. The impact model and the challenge model represent two ways of understanding how one's life could be better from a critical perspective. The model of impact measures the value of a life in terms of its consequences for the rest of the world. As Dworkin describes it,

The impact of a person's life is the difference his life makes to the objective value in the world. Impact plainly figures in our judgments about whose life was a good one. We admire the lives of Alexander Fleming and Mozart and Martin Luther King Jr., and we explain why we do by pointing to penicillin and *The Marriage of Figaro* and what King did for his race and his country. The model of impact generalizes from these examples; it holds that the ethical value of a life is parasitic on and measured by the value of its consequences for the rest of the world. The model hopes to dissipate the mysteries of ethical value by tying it to another, apparently less mysterious, kind of value: the value that objective states of affairs of the world can have. A life can have more or less value, the model claims, not because it is intrinsically more valuable to live one's life in one way or another, but because living in one way can have better consequences (251-52).

Although Dworkin thinks that the model of impact captures an important aspect of our thinking about the ethical life, it fails to recognize that much of what we value, even in our own actions and activities, is of little consequence to the rest of the world. Many people have strong ethical convictions which run counter to the model of impact:

Other people have parallel convictions: they think it important to do at least something well to master some field

of learning or craft or to learn to play a musical instrument, for example not because they will make the world better by so doing what can it matter that one more person can do something with average skill that other people can do much better? but just in order that *they* have done it. Many people set wholly adverbial goals for themselves: they want to live, they say, with integrity, doing things their way, with the courage of their convictions. These various ambitions make no sense in the vocabulary of impact. It will make no positive difference to anyone else how much or little grasp I have of cosmology, for example: I will contribute nothing to knowledge of the universe in any case. The model of impact makes many popular views about critical interests seem silly and self-indulgent (252-53).

Dworkin thinks that the impact model should be rejected due to the ways in which it leaves us with a constricted and ultimately implausible view of ethical value. In its place, he argues for the challenge model:

The alternate model that I shall now develop the model of challenge rejects [the notion that lives go better only in virtue of their impact on the objective value of states of affairs]. It adopts Aristotle's view that a good life has the inherent value of a skillful performance. So it holds that events, achievements, and experiences can have ethical value even when they have no impact beyond the life in which they occur. The idea that a skillful performance has an inherent value is perfectly familiar as a kind of value *within* lives. We admire a complex and elegant dive, for example, whose value persists after the last ripple has died, and we admire people who climbed Mount Everest because, as they said, it was there. The model of challenge holds that living a life is *itself* a performance that demands skill, that it is the most comprehensive and important challenge we face, and that our critical interests consist in the achievements, events, and experiences that mean that we have met the challenge well (253).

I take it that one of the main virtues of Dworkin's challenge model is that it offers a conception of ethical life which is more attuned to our judgments about the well-being of the agent. While the model of impact offers a familiar account of *mortality*, it is not at all clear

how the positive impact of the agent's actions is constitutive of the good or well-being of the agent. The challenge model, by contrast, with its emphasis on skillful performance and meeting the challenge of living a worthwhile life is more promising as an account of the *personal* good.

Dworkin is aware of the potential for the challenge model to be used as a justification of what he calls critical paternalism — i.e., the frustration of someone's desires for the purpose of promoting his or her critical interests. If my good consists in climbing Everest or having children or abstaining from homosexual intercourse or mastering the *Well-Tempered Clavier* regardless of whether or not I desire any of these things, then it seems that others will be justified in coercing me into doing these things, which seems decidedly the wrong conclusion. Dworkin appeals to the additive/constitutive distinction to guard against such unwanted paternalism. On the *additive view*, we can judge a person's life a good or bad one without consulting his opinions of its value. If his life has the components of a good life, then it is good for that reason. If he endorses those components, then this increases the goodness of his life; it is like frosting on the cake. But if he does not, the ethical value of the components remains. He may have a very good life in virtue of experiences and achievements he does not endorse, though not so good a life, perhaps, as if he had endorsed them (248). What Dworkin describes here as the additive view is the view which I have argued for in the course of my discussion of Parfit and Adams. He rejects the additive view in favor of the *constitutive view*, according to which no component may even so much as contribute to the value of a person's life without his endorsement. So if a misanthrope is much loved but disdains the love of others as worthless, his life is not more valuable for their affection (248). Since it is strictly impossible for something to contribute to the well-being

of an agent unless she endorses it, then critical paternalism is self-defeating.

What are we to make of Dworkin's view? Although Dworkin's concerns about the evils of paternalism are understandable and even admirable, and he employs a good deal of sophisticated analytical machinery in describing his position, Dworkin offers precious little *argument* for the endorsement constraint. This is especially problematic given the radical nature of his claim that perfectionist goods can play *no role whatsoever* in making my life better if I do not endorse them. Dworkin discusses the alleged advantages of the constitutive view in a couple of different places in *Sovereign Virtue*. In Chapter 5 (Liberal Community) Dworkin says the following:

The constitutive view is preferable for a variety of reasons. The additive view cannot explain why a good life is distinctively valuable for or to the person whose life it is. And it is implausible to think that someone can lead a better life against the grain of his most profound ethical convictions than at peace with them. If we accept the constitutive view, then we can answer the argument from critical paternalism in what we might call its crude or direct form. Suppose someone who would lead a homosexual life does not, out of fear of punishment. If he never endorses the life he leads as superior to the life he would otherwise had led, then his life has not been improved, even in the critical sense, by the paternalistic constraints he hates (217-8).

The central problem here is that, as it stands, Dworkin's claim that the additive view cannot explain why a good life is distinctively valuable for or to the person whose life it is is simply question-begging. Dworkin thinks, for instance, that a life of endorsed challenging work is better for the agent than a life of endorsed unchallenging work. Why does he think such a thing? What makes it better to endorse challenging work over easy work? Presumably, it has to do with the fact that challenging work requires the development and exercise of certain important capacities. However, even in cases where the work is not endorsed, the relevant

capacities are still developed and exercised. Thus, even unendorsed challenging work ought to contribute to the good of the agent. Dworkin's view implies that I could be a great concert pianist (the greatest the world has ever known, even) and, if I do not endorse my piano playing, perhaps because I feel that I was pressured and manipulated into this role by overdemanding parents, then my excellence qua pianist contributes absolutely nothing to my life. However, if for some reason I come to endorse my role as pianist, then suddenly my excellent piano playing – as distinguished from any pleasant feelings I might get from playing – contributes a tremendous amount to my well-being. It is hard to see how endorsement can have such transformative power.

The rest of what Dworkin says in this paragraph is true, but irrelevant to the conclusion he wants to draw. He says that it is implausible to think that someone can lead a better life against the grain of his most profound ethical convictions than at peace with them. However, we do not need to resort to the endorsement constraint to explain the truth of this claim. People who are coerced into living against their most profound ethical convictions are likely to be quite unhappy, perhaps even miserable. Profound ethical convictions are by definition the sorts of things that people care deeply about, and forcing someone to live a life in violation of them, instead of at peace with them – seems certain, or at least very likely, to impose on them a life of misery and anguish. Dworkin subtly makes use of this fact in describing the coerced homosexual as abstaining out of *fear* of punishment and as *hating* the paternalistic constraints imposed on him (218). A second distracting feature of Dworkin's rhetoric about such paternalism is the fact that the phrase "profound ethical convictions" suggests a set of values which have been arrived at via serious reflection about what constitutes a good life. It is not at all unreasonable to assume that most people who have

profound ethical convictions are likely to be, for the most part, *right* about what would make their lives better. Moreover, given the significant role that experiential states such as enjoyment and suffering will play in any plausible account of the good, along with the privileged access the agent has to her own states, we have good reason to defer to the profound ethical convictions of most people.

Dworkin offers a second discussion of the alleged superiority of the constitutive view in Chapter 6 of *Sovereign Virtue* (especially pp. 267-74). The main thing separating this discussion from the previous one is that here Dworkin associates the additive view with the model of impact and the constitutive view with the model of challenge. The impact model treats ethical value as purely objective,

...so that someone can indeed lead a better life than some alternative life he might have led, even when he thinks it a much worse life. Ethical value is additive rather than constitutive on the impact model, because ethical value is a matter of the independent value a life adds to the universe, and that cannot depend on how much value a person thinks he is adding. Creating great art does not require the artist's belief that he is creating great art. Nor does someone's improving the happiness of others require that he believe he is doing so, let alone that he believes he is leading a better life by doing so (267).

The challenge model, by contrast,

...rejects the root assumption of critical paternalism: that a person's life can be improved by forcing him into some act or abstinence he thinks valueless. Someone who accepts the challenge model might well think that religious devotion is an essential part of how human beings should respond to their place in the universe, and therefore that devotion is part of living well. But he cannot think that involuntary religious observance—prayer in the shadow of the rack—has any ethical value. He may think that an active homosexual blights his life by a failure to understand the point of sexual love. But he cannot think that a homosexual who abstains, against his own convictions and only out of fear, has

therefore overcome that defect in his life. On the challenge model, that is, it is performance that counts, not mere external result, and the right motive or sense is necessary to the right performance (269).

Dworkin's reliance on the prayer and homosexual-desire cases, with their natural focus on the interior states of agents, gives his argument here a limited application. The threat of punishment, while likely to provoke churchgoing, is unlikely to inspire true devotion, which presumably is what the religious paternalist is looking for. Similarly, if the anti-homosexual paternalist wants to change people's *desires*, then coercion is unlikely to be effective. However, if the target is instead a certain form of *behavior*, then there is no reason why such paternalism has to be self-defeating. Indeed, there are many perfectionist goods, such as athletic achievement and the mastery of difficult musical compositions, which are in principle compatible with paternalism.

It is important to note that, while Dworkin associates the model of impact with the additive view and the model of challenge with the constitutive view, there is no reason why one cannot combine the challenge model with the additive view. In other words, the additive view does not commit one to the impact model. The idea here is that part of what constitutes a good life is performing in an excellent way, or in meeting the challenge of leading an excellent life. Recall Dworkin's description of the challenge model as holding that the good life has the inherent value of a skillful performance and its expression of the idea that a skillful performance has an inherent value ...*within* lives (253). Dworkin tells us that the challenge model holds that living a life is *itself* a performance that demands skill, that it is the most comprehensive and important challenge we face, and that our critical interests consist in the achievements, events, and experiences that mean we have met the challenge well (253). With all of the emphasis here on meeting challenges and performing excellently,

it is somewhat mysterious why someone who accepts the challenge view would also commit himself to the constitutive view. Surely, if we are serious about the power of excellent performance to give value to our lives, then the additive view is the more natural account of the relation among our desires, actions, and well-being. If it is really the case that on the challenge model...it is performance that counts ¹²⁷ then performance should be enough to contribute to one's good.

So far, I have argued that Dworkin's brief for the endorsement constraint is question-begging, and given some reasons for thinking that the challenge model leads more naturally to the additive view. However, there are further reasons for rejecting the endorsement constraint as Dworkin construes it. In arguing for the endorsement constraint, Dworkin often employs the rhetorical strategy of describing two modes of life, each of which having its own advantages and disadvantages, and then asking us to imagine how an absence of endorsement would affect our estimation of the goodness (for the agent) of these modes of life. For instance, he asks us to imagine that a paternalistic government concludes that religious devotion is a waste of time, and accordingly bars people from entering religious orders (269-72). Some of these people barred from a life in the monastery turn instead to a life of politics and, even though they are quite successful, they never endorse this shift from the religious to the political life. Even if we agree with the authorities that religion is nonsense and devotion a waste of time, and that the political life is an objectively much more worthwhile endeavor, Dworkin claims that we will nevertheless think that an endorsed life in religious orders is better for an agent than a non-endorsed political life (271). However, a serious problem with examples such as this is the genuine difficulty we find in convincing ourselves that the life of

¹²⁷ See 269.

religious devotion is really worthless and that the political life is as beneficial as Dworkin suggests. Even if we think that monks have false religious beliefs and that any self-denial they engage in is pointless insofar as it depends for its justification on such false beliefs, we can still recognize ways in which life in a religious order can be beneficial. Surely, many are drawn to religious orders for the promise of a life of structure, contentment, tranquility, and a strong sense of community. By contrast, many are driven from the political life by the stress and anxiety which comes with campaigning and governing, along with the combativeness and treachery which seem such a natural part of politics. This problem is compounded by Dworkin's description of the man who endorses his life in the monastery as living with full satisfaction and confidence in his choice, whereas the man who lives an unendorsed political life finds no genuine satisfaction or self-approval and will therefore never cease regretting his choice (271). In short, such cases of endorsed religious devotion versus an unendorsed but successful political life are surrounded by too many distracting factors to allow us to make confident judgments about which is better for the agent. In order to make his point, Dworkin needs to describe a case of a life which, though endorsed, is something we can agree is definitely worthless, and then compare that to an unendorsed but objectively more worthwhile life. Only then would the case for the endorsement constraint be made. However, when we describe such a case, we are led to a very different conclusion.

Consider Scottie. Scottie had an unhappy childhood, and lives his adult years as an enraged, drug-addicted performance artist in New York (off-off-Broadway). He hates everyone, seldom bathes, never brushes his teeth, and has no real friends. However, he is devoted to his art. His art consists in coming out on stage with his remarkably untalented PunkRock back-up band, taking off his clothes, and defecating on stage. He then proceeds to

wipe his feces over his body. He then usually flings the feces at the audience. Sometimes, for special effect, he eats his feces. He gets no particular pleasure out of these performances, but they do allow him to express the rage he feels toward society. Thus, he endorses these performances; in fact, they are the central feature of his life, and virtually everything he does is for the sake of continuing with these performances. All of this culminates in his announcement that he will commit suicide on stage at the end of a particular performance. Although he is not successful in his end-of-show suicide attempt, he is able to run nude through the streets of New York to the apartment of an acquaintance, where he dies of a heroine overdose at the age of 28.¹²⁸

Now consider the life Scottie might have led. He might have become, among many other things, a moderately successful investment banker in Manhattan. Although he might find this work a bit dull, and have some ethical qualms about some of what he is expected to do in the course of his professional responsibilities, he would nevertheless lead what most of us would consider a decent enough life. This alternate Scottie has a circle of friends whose company gives him a moderate degree of pleasure, a tolerable marriage, and a minimally decent relationship with his children. He also takes advantage of the various cultural amenities offered by The City. Although Scottie leads what most of us would consider a decent enough life, he does not endorse this life. He has a vague feeling that these various social roles have somehow been thrust upon him—that the only reason he became an investment banker was pressure from his father, and that he keeps his family together mainly to please his mother. Because of this, Scottie does not endorse this life he leads, and wishes

¹²⁸ The reader may be interested to know that this hypothetical case is based on an actual person—punk rocker/performance artist G.G. Allin.

he could find the strength to leave his job and family and live the life of his performance-artist alter ego.

Which life is better for Scottie? I think that it is undeniable that Scottie-the-investment-banker is better off than Scottie-the-performance-artist. Even though he does not endorse his life as an investment banker, Scottie's life as a feces-eating performance artist is so vile, degrading, and bereft of value that it is impossible to see how anyone could recommend or approve it for a loved-one. It is a way of life one would wish only on an enemy. Of course, Dworkin, Parfit, and Adams can all agree that Scottie's life as a performance artist is worthless, since objective value is a necessary condition for something to contribute to one's well-being. Parfit and Adams, with their requirement that the agent enjoy or find pleasure in what he does, can even criticize Scottie's life insofar as it is described as being pleasureless or lacking in enjoyment. However, the problem for Dworkin is that he can make no distinction between Scottie's two possible ways of life. Both are equally worthless: the first due to its objective qualities and the second due to its not being endorsed. However, this is extremely implausible. If one had to choose one of these two lives for one's child, then it is hard to see how one could be genuinely indifferent between the two. Thus, we should reject Dworkin's claim that nothing can improve or contribute to the value of an agent's life unless it is endorsed.

We don't have to consider scenarios as extreme as Scottie's to convince us of the implausibility of Dworkin's endorsement constraint. Imagine that you live in an authoritarian regime where it is not uncommon for the government to come along and draft children for various purposes. Sometimes they take children and put them to work doing extremely menial labor; however, other times they take children and train them to be champion chess

players or great Olympic athletes. Assume that in either case the child will never endorse her life, whether it be a life of menial labor or one of competitive excellence. This might be because we think that genuine endorsement is impossible under such conditions of coercion,¹²⁹ or it could simply be that forced separation from one's parents is likely to lead to feelings of resentment which will preclude endorsement of either way of life. I take it that it makes a difference to one's quality of life whether one is coerced into being a ditch digger or a great Olympic gymnast. Even though neither is endorsed, the life of athletic excellence seems better for the agent. If one knows for certain that the state is coming for your child, it makes sense to hope that at least they will have her doing something where she can develop and exercise her capacities for excellence. Thus, *ceteris paribus*, excellence makes for a better life, but Dworkin's endorsement constraint, with its assertion that nothing can make a life better unless it is endorsed, implies that we ought to be indifferent between the coerced life of Olympic excellence and the coerced life of ditch digging. This is an implausible implication, and thus a reason to reject the endorsement constraint.

Finally, consider two different authoritarian regimes. In each, individual life plans are centrally planned. People do not decide for themselves whether they are to be doctors, lawyers, engineers, bartenders, ditch diggers, or philosophers. Under each regime these decisions are made by the Central Committee for Life Plans (CCLP). We can stipulate that in both cases the lives that people end up living are unendorsed. This could be because such coercion is incompatible with genuine endorsement, or it could be that the subjects of these regimes are so resentful of having such decisions made for them that they never endorse their ways of life. Now let's assume that the CCLP of Regime 1 has access to some very reliable

¹²⁹ See Dworkin 218.

social science which allows them to predict which individuals will flourish in which social roles: they know who are the best candidates for marriage, different kinds of jobs or professions, hobbies, and so forth. The CCLP of Regime 1 uses this scientific knowledge very conscientiously and to great effect in producing a society of successful and for most part happy people (note that feelings of satisfaction are in principle compatible with a lack of endorsement). The CCLP of Regime 2, by contrast, is both incompetent and corrupt in its planning. At best, they assign roles to people in a random fashion, so that someone with the skills and interests to be a great research scientist is made to drive a bus, while someone who would be best suited for life as a social worker is forced into computer programming. It is not hard to imagine how the quality of life in Regime 1 would be much better than that in Regime 2. However, Dworkin's endorsement constraint suggests that the two regimes are equally bad. Again, the idea that we should be indifferent between these two regimes is implausible and reason to reject the endorsement constraint.

In this section, I have argued against the claim that perfectionist goods are conditional in some way. Contrary to the claims of Parfit and Adams, one does not have to enjoy, take pleasure in, or desire a perfection in order for it to contribute to one's well-being. Dworkin's claim that something can improve or contribute to the value of someone's life only if he endorses it is an extremely strong and ultimately implausible claim. The considerations Dworkin gives in support of the endorsement constraint actually support a much weaker claim: viz., that forcing someone to act against her most profound ethical convictions is often an ineffective way of making her life better. This strikes me as undeniable. Presumably, Dworkin is attracted to the endorsement constraint because it takes paternalistic policies off the table completely. The proposition that endorsement is a necessary condition for

improving someone's life makes things much easier for the anti-paternalist. Absent the endorsement constraint, the anti-paternalist is forced into an argument about the relative value of some given excellence or perfection versus the disutility which is likely to result from the paternalistic policy. Without a theory telling us how these competing goods are to be weighed against one another, it is not completely clear how such an argument would go. However, the fact that we don't have a fully determinate theory of the personal good is a problem which cannot be remedied by accepting implausible and ultimately indefensible claims about the conditionality of perfectionist goods.

Conclusion

In this chapter, I have offered some candidates for components of the personal good. Not surprisingly, pleasure and freedom from pain, along with related experiential states such as enjoyment, contentment, anger, and depression also play a prominent role in my account. While certain states are good or bad for agents purely in virtue of how they feel, such states can also impact the agent's life in other ways. For example, grief considered simply as an experiential state is bad for the agent: i.e., it is bad *qua* pain or instance of suffering. However, considered as the appropriate emotional response to an event such as the loss of a loved one, the feeling of grief is a constituent of the agent's good: i.e., the feeling of grief *qua* emotional response benefits the agent, or contributes to his well-being. It is because we value things other than pleasure and the avoidance of pain that we seek to cultivate proper emotional responsiveness in ourselves and in those we care about. I also argued that knowledge and rationality have to be included in any plausible account of the personal good. Although the development and exercise of our capacities for theoretical and practical

deliberation can have genuine costs, it seems undeniable that we are committed to their intrinsic value. The fact that *ceteris paribus* we want those we care about (including ourselves) to exemplify rationality is evidence that knowledge and practical rationality is a constituent of the *personal* good.

I do not claim that my list of goods is necessarily exhaustive. Perhaps there are other goods which must be included as well. That is a debate for another time. My main goal here has been to identify some essential elements of the good, and explain why monistic accounts such as hedonism and perfectionism cannot possibly hope to explain or justify our most deeply held convictions about what makes a life worth living. Although the sort of value pluralism I argue for here is undeniable, it is nevertheless a controversial position among philosophers. I take it that what drives certain philosophers to value monism is the worry that pluralism precludes the rational adjudication of disagreements about value. It might be thought that what is needed in order to make judgments about how one ought to live is a simpler, more unified account of the good, or at least some clear priority rules giving us some very determinate guidance with regard to the weighing and balancing of putative goods against one another. A related issue is that of the alleged incommensurability of rival goods. It might be thought that if the various components of the good are as distinct and different from one another as my view suggests, then competing goods are simply not rationally comparable; without a common scale or supervalue to which all of the rival goods can be reduced, we have no way of making non-arbitrary judgments about what is best all things considered.

There are two issues to be addressed here. The first has to do with the plausibility of monistic accounts of the good and the prospects for a more rigorous theoretical account of the

weighing and balancing involved in all-things-considered judgments about the good. The second has to do with the costs associated with accepting pluralism. With regard to the first issue, I do not claim to have refuted every conceivable monistic account of the good, nor do I think that it is beyond the realm of possibility that someone might eventually propose a plausible set of priority rules which could give us more definite guidance in the weighing and balancing of competing goods. However, I think that our consideration of the most promising versions of value monism demonstrates that such theoretical progress is rather unlikely.

The issue of the costs associated with value pluralism is one about which I have to admit a certain degree of ambivalence. On the one hand, I am sympathetic with the desire for more theoretical unity and simplicity. I also agree that the disparateness of various types of goods makes it difficult to make publicly defensible all-things-considered judgments about value. We are forced to rely on our judgment or intuition about what seems best overall. Not only are our opponents likely to be suspicious of our appeal to judgment or intuition, but it is not even clear from our own perspective what is going on when we judge that this amount of excellence is valuable enough to outweigh that amount of pleasure. Intuitive judgments about the value of rival activities or states of affairs are significantly different from intuitive judgments about, say, the number of jelly beans in large glass jar. You and I can make conflicting intuitive judgments about the number of jelly beans, but we can see who's closer to the truth by taking out the beans and painstakingly counting them. There is no closely analogous way of testing or adjudicating conflicting intuitive judgments or opinions about the relative value of this amount of pleasure and that amount of excellence. This reliance on intuition is genuinely troubling, and my suspicion is that it is the single biggest factor among philosophers in their resistance to value pluralism.

Although a certain resistance to intuitionism is understandable, I do not think that the denial of the intrinsic value of any one of our essential goods is an intellectually respectable response to the problem. True theoretical progress is not made by arbitrarily latching on to one particular component of the good and pretending that it is the only thing which could make a life go better. While such an approach might make our all-things-considered judgments about value simpler, it would also lead people, insofar as it required them to ignore other important intrinsic goods, to live decidedly impoverished lives. Another problem with the intuitionism-is-bad objection is that it fails to recognize that even monistic accounts will require the use of, and appeal to, the faculty of judgment or intuition. For instance, even if we accept a quality-of-experience view such as hedonism we will have to make intuitive judgments about the relative value of competing types of experiences or types of pleasure or pain. For, as we have seen, pleasures form a decidedly heterogeneous class, and there is apparently nothing they have in common other than intrinsic desirability; moreover, we must not lose sight of the fact that pleasure and freedom from pain seem to constitute two distinct, heterogeneous states. If we accept Hurka's version of perfectionism, we will have to make intuitive judgments about the relative value of instances of theoretical perfection, practical perfection, and physical perfection, and all of the indifference curves in the world cannot mask the fact that we simply do not have any mechanical, non-intuitive methods for making such judgments. The version of perfectionism championed by Green and Brink seems no better positioned to avoid the appeal to intuition in determining what would constitute the proper exercise of our capacities for rational deliberation. Thus, even if we accept a version of value monism, this will not save us from the uncertainty and embarrassment of appeal to intuition.

Sometimes it seems that the worry about pluralism is a more straightforwardly *practical* concern. How can we even make judgments about what is best without a rigorous, monistic theory? This worry, insofar as it is distinguished from the worry above about the ultimate epistemic status of our judgments about value, is hard to take seriously. People have been making judgments about what is best all things considered from time immemorial. There is nothing new about it. The real threat is that we will lose sight of the complexity of the good and develop an obsession with some particular aspect of the good, thereby living impoverished lives. This can happen when we lose sight of our own complexity as embodied rational agents. Thus, the hedonist forgets that we are not mere animals but also rational agents, and recommends a life more fit for swine than men, while the perfectionist ignores the significant similarities between human beings and the rest of the animal kingdom and recommends a life better suited to disembodied rational agents. Hedonism and perfectionism both have a tendency to ignore the emotional aspects of our nature, and thus typically fail to accord proper emotional responsiveness the position it deserves in a plausible account of the good. A proper understanding and appreciation of our own complex natures allows us to identify those things which can make our lives better, and can help us to make judgments about the relative value of competing goods. Thus, the relative value we attribute to knowledge or rationality will depend on the importance we place on our status as rational agents. For instance, given the role of rationality in explaining why human beings have a special moral status or dignity deserving of respect, the development and exercise of our capacities for rational deliberation should rank relatively high with regard to value.

In order to determine what would be best for a particular person, we can't just think about the relative value of knowledge, practical rationality, emotional responsiveness, and

pleasure. We also have to give serious consideration to facts about that individual person. His well-being, I take it, is a function of the value of each component of the good in his life. Any intrinsic good which a person experiences or exemplifies – be it pleasure, emotional responsiveness, knowledge and rationality, and so forth – contributes to his well-being. However, some people will be better candidates for certain life plans than others. Someone who is extremely adept at abstract thought but possesses poor interpersonal skills is much more likely to flourish (both in terms of excellence and enjoyment) as an engineer than as a police officer or teacher. Someone with a great natural singing voice but poor manual dexterity would be much better off pursuing a career in the opera than one as a sculptor or painter. Someone with an unusually strong disposition toward feelings of stress and anxiety when put in demanding situations might very well be better off working in a surf shop than pursuing an especially challenging profession. Someone who is predisposed to a certain level of contentment regardless of her activities should perhaps gravitate toward especially demanding work where she could achieve various perfectionist goods and thus improve her life in that way.

Our judgments about what would be best for a person are not as difficult as pluralism might suggest at first glance. As a matter of fact, people do tend to enjoy doing that at which they excel. Thus, we should be careful not to exaggerate the conflict or tension between pleasure and excellence. Also, facts about my own personal psychology and situation can help determine what sort of life I would be best off pursuing. The amount of good I can achieve in various courses of life will depend on my own abilities, dispositions, and resources. Thus the familiar phenomenon of different people being better suited for radically different forms of life. Of course, there are things we can do to change an individual's

situation. For example, changes in the environment can give rise to radically different basic abilities and dispositions: it makes a tremendous difference to my abilities and dispositions whether I am raised in impoverished Eastern Kentucky or the affluent Upper West Side of Manhattan. The changes we make to social environments will depend on our judgments about the relative value of the ways of life open to products of these environments. Presumably, the life of the affluent Manhattanite will score higher along most of our dimensions (though there is no guarantee of this); thus, we would probably be fairly confident that, in moving the infant Sallie from Kentucky to Manhattan, we would be giving her a better life all things considered.

The progress that is being made in various kinds of genetic technologies gives us reason to think that we will be able in the not-too-distant future to make changes to the genetically-based abilities and dispositions of individuals. This has the promise of removing certain facts about the individual as constraints on what kind of life it is sensible for her to pursue; for the abilities and dispositions we would point to will become increasingly malleable. If such genetic manipulation becomes possible and accepted, then this will make our decisions about what is best for someone increasingly difficult. For then we will be put in the position of having to decide what would be best simply in terms of the relative value of the various intrinsic goods. We will have to ask, What is better, knowledge or enjoyment? Obviously, such questions are hard to answer. We can only hope that, if such possibilities do come to pass, people are aware of the difficulty in making judgments about what constitutes a good life, and that they are hesitant to impose their own conception of a good life on others through genetic manipulation in the same way that we are currently wary of manipulating people's lives through radical forms of social engineering. The prospects of such genetic

manipulation makes it even more crucial that those with the power to alter an individual's nature be sensitive to the complexities of the good and the various ways in which a life can be made better or worse.

Bibliography

- Adams, Robert. 1999. *Finite and Infinite Goods*. Oxford: Oxford University Press.
- Aristotle. *Nicomachean Ethics*. 2nd edition. Translated and edited by Terence Irwin. Indianapolis: Hackett Publishing Company, 1999.
- Arneson, Richard J. 1999. Human Flourishing versus Desire Satisfaction. In *Human Flourishing*, edited by Ellen Frankel Paul, Fred D. Miller Jr., and Jeffrey Paul. Cambridge: Cambridge University Press.
- Barry, Brian. 1965. *Political Argument*. London: Routledge and Kegan Paul.
- Beauchamp, Tom and James Childress. 2001. *The Principles of Biomedical Ethics*, 5th edition. Oxford: Oxford University Press.
- Bentham, Jeremy. 1823. *Introduction to the Principles of Morals and Legislation*. London: Athlone, 1970.
- Boyd, Richard. 1988. How to Be a Moral Realist. In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord. Ithaca: Cornell University Press.
- Bradley, F.H. 1927. *Ethical Studies*. 2nd edition. Oxford: Clarendon Press.
- Brandt, Richard. 1979. *A Theory of the Good and the Right*. New York: Oxford University Press.
- Brink, David. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Brink, David. 1992. Mill's Deliberative Utilitarianism. *Philosophy and Public Affairs* 21:67-103.
- Brink, David. 1997a. Rational Egoism and the Separateness of Persons. In *Reading Parfit*, edited by Jonathan Dancy. Oxford: Blackwell Publishers.
- Brink, David. 1997b. Self Love and Altruism. *Social Philosophy and Policy* 1:122-57.
- Brink, David. 2003. *Perfectionism and the Common Good: Themes in the Philosophy of T.H. Green*. Oxford: Oxford University Press.
- Broad, C.D. 1930. *Five Types of Ethical Theory*. London: Routledge and Kegan Paul.

- Carson, Thomas. 1995. Review of *Perfectionism*, by Thomas Hurka. *Philosophy and Phenomenological Research* 55: 719-23.
- Daniels, Norman. 1979. Wide Reflective Equilibrium and Theory Acceptance in Ethics. *Journal of Philosophy* 76: 256-82.
- Daniels, Norman. 1980. Reflective Equilibrium and Archimedean Points. *Canadian Journal of Philosophy* 10: 83-103.
- Dupré, John. 1981. Natural Kinds and Biological Taxa. *Philosophical Review* 90: 66-90.
- Dworkin, Ronald. 2000. *Sovereign Virtue*. Cambridge, Mass: Harvard University Press.
- Finnis, John. 1980. *Natural Law and Natural Rights*. Oxford: Oxford University Press.
- Frankena, William. 1973. *Ethics*. 2nd edition. Englewood Cliffs, NJ: Prentice-Hall.
- Gaut, Berys. 2002. Justifying Moral Pluralism. In *Ethical Intuitionism: Re-evaluations*, edited by Philip Stratton-Lake. Oxford: Clarendon Press.
- Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Cambridge, Mass.: Harvard University Press.
- Gosling, J.C.B. 1969. *Pleasure and Desire: The Case for Hedonism Reviewed*. Oxford: Oxford University Press.
- Green, T.H. 1883. *Prolegomena to Ethics*. Edited by David Brink. Oxford: Oxford University Press, 2003.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press.
- Hardie, W.F.R. 1968. *Aristotle's Ethical Theory*. Oxford: Oxford University Press.
- Harman, Gilbert. 1982. Critical Review: Richard B. Brandt, *A Theory of the Good and the Right*. *Philosophical Studies* 42, 119-39.
- Hurka, Thomas. 1987. The Well-Rounded Life. *Journal of Philosophy* 84: 707-26.
- Hurka, Thomas. 1993. *Perfectionism*. New York: Oxford University Press.

- Hurka, Thomas. 1999. The Three Faces of Flourishing. *Social Philosophy and Policy* 16: 44-71.
- Hurka, Thomas. 2001. *Virtue, Vice, and Value*. Oxford: Oxford University Press.
- Hylton, Peter. 1990. *Russell, Idealism, and the Emergence of Analytic Philosophy*. Oxford: Clarendon Press).
- Johnston, Mark. 1987. Human Beings. *Journal of Philosophy* 84: 59-83.
- Kant, Immanuel. 1785. *Groundwork of the Metaphysics of Morals*. Translated and edited by Mary Gregor. Cambridge: Cambridge University Press, 1997.
- Kitcher, Philip. 1999. Essence and Perfection. *Ethics* 110: 59-83.
- Kraut, Richard. 1979. Two Conceptions of Happiness. *Philosophical Review* 88: 176-96.
- Kraut, Richard. 1994. Desire and the Human Good. *Proceedings and Addresses of the American Philosophical Association* 68: 39-54.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
- Kymlicka, Will. 1989. *Liberalism, Community, and Culture*. Oxford: Oxford University Press.
- Lacey, Alan. 1995. Naturalism. In *The Oxford Companion to Philosophy*, edited by Ted Honderich. Oxford: Oxford University Press.
- LaPorte, Joseph. 1997. Essential Membership. *Philosophy of Science* 64: 96-112.
- Lawrence, Gavin. 1993. Aristotle and the Ideal Life. *Philosophical Review* 102: 1-35.
- Locke, John. 1700. *An Essay Concerning Human Understanding*. Edited by P.H. Niddich. Oxford: Clarendon Press, 1975.
- Long, A.A. and Sedley, D.N. 1987. *The Hellenistic Philosophers Vol. 1*. Cambridge: Cambridge University Press.
- Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin Books.

- Mill, John Stuart. 1861. *Utilitarianism*. Edited by George Sher. Indianapolis: Hackett Publishing Company, 1979.
- Moore, G.E. 1903a. *Principia Ethica*. Cambridge: Cambridge University Press.
- Moore, G.E. 1903b. The Refutation of Idealism. *Mind* 12: 433-53.
- Mulgan, Tim. 1994. Review of *Perfectionism*, by Thomas Hurka. *Mind* 103: 550-53.
- Murphy, Mark. 2001. *Natural Law and Practical Rationality*. Cambridge: Cambridge University Press.
- Nagel, Thomas. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.
- Nicholson, Peter. 1990. *The Political Philosophy of the British Idealists*. Cambridge: Cambridge University Press.
- Noakes, J. and G. Pridham. 1988. *Nazism: A History in Documents and Eyewitness Accounts, 1919-1945*, Vol. 2. New York: Schocken Books.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Post, John. 1995. Naturalism. In *The Cambridge Dictionary of Philosophy*, edited by Robert Audi. Cambridge: Cambridge University Press.
- Putnam, Hilary. 1970. Is Semantics Possible? Reprinted in Putnam 1975b.
- Putnam, Hilary. 1975a. The Meaning of Meaning . Reprinted in Putnam 1975b.
- Putnam, Hilary. 1975b. *Mind, Language and Reality: Philosophical Papers, Volume 1*. Cambridge: Cambridge University Press.
- Railton, Peter. 1986a. Facts and Values. *Philosophical Topics* 14: 5-31.
- Railton, Peter. 1986b. Moral Realism. *Philosophical Review* 95: 163-207.
- Rawls, John. 1951. Outline of a Decision Procedure for Ethics. *Philosophical Review* 60: 177-97.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.

- Rawls, John. 1974. The Independence of Moral Theory. *Proceedings and Addresses of the American Philosophical Association* 48: 5-22.
- Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.
- Rosati, Connie. 1995. Persons, Perspectives, and Full Information Accounts of the Good. *Ethics* 105: 296-325.
- Ross, W.D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- Ross, W.D. 1939. *The Foundations of Ethics*. Oxford: Clarendon Press.
- Salinger, J.D. 1945. *The Catcher in the Rye*. Boston: Little, Brown and Company.
- Schmitt, Frederick. 1995. Naturalism. In *A Companion to Metaphysics*, edited by Jaegwon Kim and Ernest Sosa. Oxford: Blackwell Publishers.
- Sher, George. 1997. *Beyond Neutrality*. Cambridge: Cambridge University Press.
- Sidgwick, Henry. 1877. Critical Notice: *Ethical Studies* by F.H. Bradley. *Mind* 2:125-6.
- Sidgwick, Henry. 1907. *The Methods of Ethics*. 7th edition. Indianapolis: Hackett Publishing Company, 1981.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell Publishers.
- Sobel, David. 1994. Full Information Accounts of Well-Being. *Ethics* 104: 784-810.
- Sober, Elliot. 1980. Evolution, Population Thinking, and Essentialism. *Philosophy of Science* 47: 350-83.
- Sober, Elliot. 1993. *Philosophy of Biology*. Boulder: Westview Press.
- Sprigge, T.L.S. 1983. *The Vindication of Absolute Idealism*. Edinburgh: Edinburgh University Press.
- Stratton-Lake, Phillip. 2002. Pleasure and Reflection in Ross's Intuitionism. In *Ethical Intuitionism: Re-evaluations*, edited by Phillip Stratton-Lake. Oxford: Clarendon Press.
- Sumner, L.W. 1995. Review of *Perfectionism*, by Thomas Hurka. *Philosophical*

Review 104: 151-53.

Sumner, L.W. 1996. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press.

Velleman, J. David. 1988. Brandt's Definition of Good. *Philosophical Review* 97: 353-371.

Whiting, Jennifer. 1988. Aristotle's Function Argument: A Defense. *Ancient Philosophy* 8: 33-48.

Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, Mass.: Harvard University Press.

Williams, Bernard. 1993. *Morality: An Introduction to Ethics*. Cambridge: Cambridge University Press.

Wolfe, Tom. 2004. *I am Charlotte Simmons*. New York: Farrar, Straus and Giroux.

Wollheim, Richard. 1959. *F.H. Bradley*. Harmondsworth: Penguin Books.