

**The Reward of Virtue: An Essay on the
Relationship Between Character and Well-Being**

**A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY**

Ian M Stoner

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

Valerie Tiberius and Michelle Mason, Advisors

July, 2011

© Ian M Stoner 2011

Acknowledgments

Thanks to professors Fred Schueler, Doug Lewis, Sandra Peterson, Michelle Mason, and Valerie Tiberius for leading first-rate seminars that planted the seeds that eventually became this dissertation. Thanks to the members of several writing groups along the way, including at least Marilea Bramer, Mike Rohde, Matt Frank, Jason Swartwood, and Steve Nelson. Thanks to Devora Shapiro for discussions during the formative stages of these ideas. Thanks to Crystal Bergstrom for years of patience followed by a frantic marathon of copy-editing.

Finally, thanks to Valerie Tiberius for embodying the virtuous advisor.

Dedication

For my parents, Mark and Daria, with gratitude for decades of unflagging support.

Abstract

Most work in neo-Aristotelian virtue ethics begins by supposing that the virtues are the traits of character that make us good people. Secondary questions, then, include whether, why, and in what ways the virtues are good *for* the people who have them.

This essay is an argument that the neo-Aristotelian approach is upside down. If, instead, we begin by asking what collection of character traits are good for us—that is, what collection of traits are most likely to promote our own well-being—we find a collection of traits a lot like the traditional slate of virtues.

This suggests an egoistic theory of the virtues: the virtues *just are* those traits of character that reliably promote the well-being of their possessor. In addition to making the positive case for character egoism, I defend it from some anticipated objections. Most importantly, I argue that character egoism doesn't inherit the problems of ethical egoism. I conclude by offering self-regarding accounts of two virtues traditionally thought to be irreducibly other-regarding: honesty and justice.

Contents

Acknowledgments	i
Dedication	ii
Abstract	iii
1 Introduction: Virtue and Well-Being	1
1.1 Purpose, function, and excellence of kind	4
1.2 The connection between virtue and well-being	8
1.3 Where from here?	11
2 Doing Well Without a Theory of Well-Being	12
2.1 Relationships	14
2.2 Personal projects	17
2.3 Other candidates	21
2.4 Projects, relationships, and theories of well-being	27
2.5 Character egoism	30
3 Ethical Egoism and Character Egoism	34
3.1 Egoism as category mistake	36
3.2 Egoism and beneficence	38

3.3	Egoism and moral schizophrenia	43
3.4	Conclusion	54
4	Clearing Ground for Egoistic Virtues	56
4.1	“But, Hitler!” and the threat of endorsing vices	57
4.2	Bad projects and the threat of circularity	62
4.3	Context-sensitivity and the threat of instability	70
4.4	The situationist challenge to virtue ethics	76
5	A Catalog Grounded in Well-Being	83
5.1	Getting a catalog out of the definition	83
5.2	A fragment of a catalog of virtues	86
5.3	Conclusion	107
6	Egoistic Honesty: Responding to the Sensible Knave	108
6.1	Hume’s problem	110
6.2	Trust, friendship, and knavery	116
6.3	Conclusion	124
7	Character Egoism and Social Justice	125
7.1	Oppressed groups and the pressure to adapt	126
7.2	Social justice as a self-regarding virtue	135
8	Conclusion	157
	References	160

Chapter 1

Introduction: Virtue and Well-Being

But where is the reward of virtue? And what recompence has nature provided for such important sacrifices, as those of life and fortune, which we must often make to it?

“The Stoic”

DAVID HUME

Consider the two morally best people you know. They aren't, presumably, clones. They differ in superficial ways: maybe one is tall and the other short. They have different personalities, too: one might have a sharp sense of humor while the other is outgoing.

They also differ from one another in ways that are clearly morally relevant. Perhaps one is kinder while the other is more steadfast. We can account for the morally relevant differences between these people in two ways. We could insist that every morally relevant difference between them represents an aspect of their character in which at least one of them is morally *deficient*. Alternatively, we

might suppose that these two people are comparably good *in different ways*.

In fact, the wide variety of human life-roles yields a ready supply of examples in which we judge different moral characters as comparably good in different ways. A good poet should cultivate a degree of emotional sensitivity that would be crippling (and lamentable) for a military commander. A diplomat should cultivate dispositions that allow her to move comfortably between cultures, while a farmer with no interest in travel need not. The poet and the commander, the diplomat and the farmer, are good in different ways.

While it is easiest to see different people being good in different ways when those people have different life-roles, the differences in their characters can't be chalked up (at least entirely) to their different life-roles. Consider, for example, the wide range of characters that could rightly be described as excellent (or virtuous) teachers. Some excellent teachers are stern and demanding. Some excellent teachers are nurturing and emotionally connected with their students. Some are passionate and funny. They are good teachers, in different ways. And when a situation with clear moral stakes arises in the classroom—say, a student says something patently bigoted during a discussion—excellent teachers will deal with it in different ways. They will be virtuous in different ways.

There is variety among the virtuous.

At the same time, the degree of overlap in the characteristics of virtuous people is striking. When you think of the two best people you know, it may be the case that one is kinder than the other, but it will never be the case that either of the best people you know is *cruel*. The traits of character that make a person virtuous are broadly shared.

Neo-Aristotelian virtue theories, which ground the virtues in *human* nature or excellence of *kind*, are well-positioned to account for the similarities between

virtuous people. Human nature, after all, is shared between poets, generals, teachers, and everyone else. The slate of virtues derived from human nature will be common to us all.

For the very same reason, though, theories in the Aristotelian tradition are less well-positioned to account for individual variation in the characters of virtuous people. There are many kinds of variation neo-Aristotelians can capture, but given that human nature is shared between us all, different virtues for different people appears to be incompatible with neo-Aristotelian theory.

We have available to us another concept that can ground an account of the virtues: well-being. Well-being is a feature of individuals, not species, so an account of virtue grounded in well-being will naturally accommodate variety in the characters of the virtuous. However, it is a psychological fact that the sorts of engagement that foster well-being are broadly shared across the kind. For all human beings, close friendships are good; for all human beings, a sense of achievement is good. The fact that the psychology of well-being is broadly shared accounts for the fact that well-being derived virtues, too, are broadly shared.

I believe this is the better way to approach theorizing about the virtues. The virtues are best understood as the traits of character that reliably promote our own well-being; that is, the virtues are best understood *egoistically*. Though the poet, the general, the farmer, the diplomat, and the teacher are all instances of the human kind, they have reason to cultivate varying slates of virtues: their lives will go better *for them* if they do.

Neo-Aristotelian virtue ethics and its thirty years of off-shoots have together developed a framework that includes much more than a definition of virtue. Most centrally, the various virtue theories together make a compelling case that character is (or should be) prior to action in moral evaluation; indeed, they all derive

an account of right action from an account of good character. This understanding of the relationship between character and action is an inversion of the approach of modern moral theories like utilitarianism and Kantianism, and I believe it is right.¹

My disagreement, then, is not with the general framework of virtue ethics. My concern is specifically with the neo-Aristotelian theory of the virtues as the traits of character that make us excellent instances of our kind.² This dissertation is not a direct argument against the Aristotelian theory of the virtues; it is, instead, an alternative theory. By way of introduction, however, I'd like to run through some well-trod criticisms of neo-Aristotelian virtue theory. In reviewing these issues, I'm not offering anything new—let alone a knockdown argument against neo-Aristotelians. I want only to draw attention to the difficulty (and openness) of some questions in the neo-Aristotelian tradition, and to the ways in which these questions invite us to reconceive the connection between virtue and well-being.

1.1 Purpose, function, and excellence of kind

Virtue theories in the Aristotelian tradition take the virtues to be the characteristics that make a thing an excellent instance of its kind. Being an excellent instance of one's kind consists in *functioning* well, or serving one's *purpose* well.

Consider a knife. The purpose of a knife is to slice things. To function well in this role, a knife needs a strong handle and a sharp blade. These—strong handle, sharp blade—are the *virtues* of a knife. They are the characteristics that make a knife a *good* knife.

¹ See Oakley 1996 for a discussion of the areas of agreement and disagreement among various virtue theories.

² Or, alternatively, as the traits of character that contribute to *eudaimonia*, where eudaimonia is well-being's normatively loaded cousin.

Aristotle offers other examples:

Every virtue causes its possessors to be in a good state and to perform their functions well; the virtue of the eyes, e.g., makes the eyes and their functioning excellent, because it makes us see well; and similarly, the virtue of a horse makes the horse excellent, and thereby good at galloping, at carrying its rider and at standing steady in the face of the enemy. If this is true in every case, then the virtue of a human being will likewise be the state that makes a human being good and makes him perform his function well (NE 1106a14ff).

Neo-Aristotelian virtue ethicists like Hursthouse and Foot take the same basic approach: just as the virtues of a knife are the characteristics that make it a good knife, the virtues of a person are the traits of character that make her a good person.

But, it turns out, adapting Aristotle's approach for the modern world requires philosophical heavy lifting.

We give the knife its purpose—we designed it to cut things. *We* give the warhorse its purpose—we bred it to be big and trained it to be fearless. But this notion of purpose is inappropriate for a general account of virtue or goodness. There are many things in the natural world that exist without being given a purpose by people. There are oak trees, for example, that no one has ever seen, let alone cared to assign a purpose. Oak trees, lacking the ability to plan and have intentions, can't give *themselves* purposes. Absent the natural teleology that organized the world for Aristotle, we must reinterpret the idea of purpose, at least a little bit, if we are to make sense of the idea of excellence of kind. Reinterpretation turns out to be hard to do.

In *Natural Goodness*, Philippa Foot tries to preserve something close to Aristotle in describing a causal or functional teleology she thinks should be acceptable to moderns. She connects virtue (or natural goodness) to “the relation of an individual to the ‘life form’ of its species” (Foot 2003, 27). There are statements about the world (which she calls “Aristotelian categoricals”) that capture something like teleological facts about individuals and species. For example, “there is an Aristotelian categorical about the species *peacock* to the effect that the male peacock displays his brilliant tail *in order to* attract a female... The display serves this purpose.” In general, “what is crucial to all teleological propositions [i.e. Aristotelian categoricals] is the expectation of an answer to the question ‘What part does it play in the life cycle of things of the species?’ In other words, ‘What is its function?’ or ‘What good does it do?’” (Foot 2003, 31-32). Thus it is that we can give evaluations of things without appeal to our own values; thus it is there can be *natural* goodness. First we understand the relevant Aristotelian categoricals—the facts that correspond to the teleology of the species—and then we evaluate how well the thing in question matches up with its categorical.

She gives, as example, an evaluation of an oak tree:

We are, let us suppose, evaluating the roots of a particular oak tree, saying perhaps that it has good roots because they are as sturdy and deep as an oak’s roots should be. Had its roots been spindly and all near the surface they would have been bad roots; but as it is they are good. Oak trees need to stay upright because, unlike creeping plants, they have no possibility of life on the ground, and they are tall heavy trees. Therefore oaks need to have deep sturdy roots: there is something wrong with them if they do not, and this is how the normative proposition can be derived. The good of the oak is its

individual and reproductive life cycle, and what is necessary for this is an Aristotelian necessity in its case. Since it cannot bend like a reed in the wind, an oak that is as an oak should be is one that has deep and sturdy roots (Foot 2003, 47).

Foot is clear that she doesn't intend her causal or functional natural teleology to boil down to evolutionary biology (32). It isn't so clear, though, that she can avoid this result.

If she abandons strong talk of purposes (that is, purposes assigned to things by people) in favor of functional or causal accounts of things, then it seems that she's abandoned talk of teleology entirely. The *goal* of the oak tree's leaves is not to provide shade for the picnickers beneath. The *goal* of the oak tree's deep roots is not to withstand the thunderstorm forecast to roll through on Wednesday. The *goal* of an oak tree's phloem is not to carry nutrients from the roots to the leaves. The shade, the sturdiness, the capillary action, are merely the effects of the oak tree's structure. That's it. Absent some standard, like fitness or purpose or fitness to reproduce, no account of the structure of an oak tree, or of the role its parts play in causal chains, is enough to support the *normative* judgments—the *evaluations*—Foot wants.

Despite her desire to avoid it, evolutionary biology *can* provide an external standard by which the systems of an oak tree can be judged: reproductive fitness. Sturdy roots have been selected, over millions of years, because they tend to keep the tree alive during storms. Phloem has been naturally selected because it is effective at carrying nutrients from the roots to the leaves.

I suppose Foot wants to avoid the evolutionary account of function because the sorts of evaluations it underwrites aren't much like the evaluations of natural goodness she's after. Consider a mad villain who poisons all the world's

groundwater. Suddenly, oak trees with deep and sturdy roots will die, leaving as the only reproductively successful trees those with superficial (and thus un-oaky) roots. Which are the naturally good (that is, Footian) oak trees, in this case?

These sorts of radical changes of environment have happened and happened again in the evolutionary history of every species. This highlights the fact that evolution talk isn't uncovering any kind of goal-directedness in oak trees. Evolutionary evaluations merely assign goodness to one sort of outcome—reproduction—from a set of causal interactions that isn't goal-directed in any way.

It's even harder to connect this evolutionary account of goodness to human beings in a way that isn't offensive or silly. While it might make sense to judge the excellence of an oak tree by the number of viable offspring its features would tend to produce, it makes no sense to judge people that way.

Though I've focused on Foot, the problem of adapting Aristotle for a world that no longer believes in biological teleology, or a purpose-driven natural world, is a challenge for all theories that begin with the idea of humanity or a human kind. In order to embrace the Aristotelian approach, we must first develop an account of what it means to be an excellent instance of the human kind, and this isn't easy to do.³

1.2 The connection between virtue and well-being

Another problem for theories in the Aristotelian tradition: if our virtues are what make us an excellent instance of our kind, then the connection between virtue and well-being is mysterious. And yet most Aristotelians—most people, for that

³ Thanks to G.F. Schueler for discussions of teleology in Foot's book.

matter—share a conviction that virtue and well-being strongly correlate. Rosalind Hursthouse, for example, identifies three theses that characterise Aristotelian virtue ethics: “1. The virtues benefit their possessor... 2. The virtues make their possessor a good human being... 3. The above two features of the virtues are interrelated” (Hursthouse 2002, 167). Explaining this connection occupies the last third of her *On Virtue Ethics*.

The possibility of a gap between virtue and well-being is clearly present in the strong “purpose” reading of Aristotle. Peter Glassen long ago argued that to suppose the human good is rational activity in accordance with virtue is a “confusion of the notion of *the goodness of* with the notion of *the good of*” (Glassen 1957, 319). Or, as Korsgaard elaborates Glassen’s point, “Might not a skittish unmanageable horse win for itself a fine free horse-life away from the dangers of warfare?” (Korsgaard 2008, 131). A bodyguard who serves her purpose—absorbing bullets—does not thereby secure her own good. The canary that serves its purpose by dying at the first whiff of methane does not thereby secure its own good. There is no necessary connection between fitness to purpose and well-being.

The causal/evolutionary approach is no more promising, at least if we suppose—as I do—that well-being has an affective component. (It usually *feels good* to be doing well.) Why should evolution have made us *enjoy* or feel *satisfied by* being an excellent instance of our kind? Excellent rabbits, after all, apparently spend their lives in abject terror; excellent oak trees don’t feel much at all.

These, then, are the difficult problems posed by the idea of species-excellence that neo-Aristotelians inherit from Aristotle.⁴ Much work has been done,

⁴ These issues come to a head in the “function argument” passage at *Nicomachean Ethics* 1097b-1098a. Critics have complained about nearly every aspect of the function argument. Christine Korsgaard rounds some of them up:

The idea that human beings even have a function is supposed to be based on a dubious teleological principle or an illegitimate piece of teleological reasoning. The

and continues to be done, in solving these problems. Everyone working in the neo-Aristotelian tradition is aware of the criticisms of the function argument, and none of them thinks the situation is dire. I don't mean to suggest otherwise. However, thinking about the complications that follow on grounding ethics in human function or excellence of kind suggests an alternative approach that sidesteps entirely these problematic notions. Instead of understanding the virtues as traits that serve our characteristic *human excellence*, we should understand the virtues as traits that serve our individual *well-being*.

Human well-being is itself a complicated and controversial concept. Nevertheless, while it isn't clear that there *is* such a thing as the human function, we can be confident that whatever it is, there is such a thing as well-being. More, characterizing the virtues as traits that promote well-being builds in a clear connection between virtue and well-being.

This dissertation is an argument that well-being provides a solid foundation for a theory of the virtues. That is, I argue that a *character-egoist* theory of the virtues accounts for our moral intuitions at least as well as does neo-Aristotelian virtue theory, and it does it with less theoretical baggage.

inference that the good performance of this function, supposing that it did make you a good human being, would therefore be good for you, has been deemed a "fallacy." The assumption that the good performance of the function would make you a good human being is called into question by the thought that any human capacity can be used—and used, in a non-moral sense, excellently—either for good or for evil. Even if these problems were resolved, Aristotle's method of selecting the function—by choosing the kind of life that is unique to human beings—raises a whole new set of problems, since his critics cannot see either why it should be one of these or why it should be the one that is unique (Korsgaard 2008, 133).

The other problems are, in my view, less fundamental and more tractable than the problems of purpose and the relationship between virtue and well-being.

1.3 Where from here?

Most of this dissertation is devoted to clearing three major hurdles from the path of character egoism. The first hurdle is the worry that the concept of individual well-being is every bit as controversial and poorly understood as the concept of human excellence. In chapter two, “Doing Well Without a Theory of Well-Being,” I argue that we don’t need a theoretical account of well-being to develop a normative theory grounded in well-being. A minimal, “practical” account of well-being will do. I sketch this practical account, and show that it is compatible with most going theories of well-being, both in philosophy and psychology.

The second hurdle is the nearly universal dismissal of any moral theory with “egoism” in the name. In chapter three, “Ethical Egoism and Character Egoism,” I argue that much of the horror of egoism is rooted in serious philosophical problems with *ethical* egoism, but these problems do not infect character egoism. Character egoism should be treated as a distinct kind of theory and evaluated on its merits, not dismissed because it shares a family name with ethical egoism.

The third hurdle is the belief that no egoistic approach can hope to accommodate the bedrock virtues of honesty and justice, both of which are usually thought to be irreducibly other-regarding virtues. In chapters six and seven, I give self-regarding accounts of honesty and justice.

In the course of these mainly defensive chapters, I offer the beginnings of a positive theory of character egoism. In chapter four, “Clearing Ground for Egoistic Virtues,” I sketch the philosophical and psychological limits in which such a theory must develop. In chapter five I give several examples of a character-egoist method of cataloging virtues and vices.

Chapter 2

Doing Well Without a Theory of Well-Being

Happiness cannot be pursued; it must ensue, and it only does so as the unintended side effect of one's personal dedication to a cause greater than oneself or as the by-product of one's surrender to a person other than oneself.

Man's Search for Meaning

VICTOR FRANKL

Whatever well-being is, it is a feature not of species, but of individuals. Well-being, as L.W. Sumner puts it, “has to do with what we may call the *prudential value* of a life, namely, how well it is going *for the individual whose life it is*” (Sumner 1998, 21).

Though giving a full account of, or *defining* well-being is a book-length project of its own, arm-chair philosophizing and psychological research together paint a fairly consistent picture of what sorts of things we need to do if we are to live our

lives well from our own point of view. According to philosophers and psychologists alike, we do well in our lives when we do well in the projects and relationships that are of special interest or importance to us.

A person's set of projects and relationships is special (particular, unique) to him or her. There is no kind of project (say, oboe playing) or kind of relationship (say, parent) that will promote the well-being of every person. One must know a lot about a person's values, interests, strengths and weaknesses, circumstances and potential opportunities, before it is possible to guess at what projects and relationships are the sort that would probably be important to, and thus satisfying for, her.

Success in projects and relationships about which I care little is little satisfying. If I find myself a sought-after theater director, while at the same time loathing everything about the theater, my life is not going well.¹ Similarly, if I am failing at projects and relationships that *are* important to me, my life is not going well. If I love and value my spouse and yet my marriage is crumbling, my life is not going well. In short: if I'm not doing well in the projects and relationships that matter to me, then I'm not doing well.

Conversely, if the projects and relationships that are important to me are going well, there is little that could block the judgment that I am doing well. If I'm very poor, but my poverty doesn't inhibit my projects or stress my relationships, I'm doing well in the face of adversity. If I have a chronic disease, but do well in my relationships and projects despite, I'm doing well in the face of adversity.

In this chapter I argue that relationships and projects together constitute a uncontroversial *proxy* for well-being. It is true that there is no consensus about which is the best theory of well-being. For the purposes of a normative virtue

¹ I have in mind Darren Nichols, the pseudo-Brechtian theater director on *Slings and Arrows*.

theory, that doesn't matter. We have, in projects and relationships, a proxy that's entirely adequate to support a theory of the virtues that's grounded in well-being.²

2.1 Relationships

The view that friendships and other relationships are central to human well-being is well-represented in the history of philosophy. Aristotle says, “without friends, no one would choose to live, though he had all other goods” (NE 1155a5).

In *Laelius: On Friendship*, Cicero writes:

The benefits that friendship offers almost defy description, they are so great. To begin with, how can life be “worth living” at all... unless it reposes on the mutual goodwill of friends? It is the most satisfying experience in the world to have someone you can speak to as freely as your own self about any and every subject on earth (188).

Montaigne's “On Friendship” is a long meditation on how his friendship with Etienne de La Boétie enriched his life.

² Someone (someone like my friend and grad-school colleague Matt Frank) might ask why I treat relationships and projects as separate categories, rather than giving an account of one in terms of the other, or in terms of an underlying value they have in common.

It is true that relationships and projects are inevitably tangled up with each other. (For example, the relationship of mentor and the project of mentoring are difficult or impossible to separate.) Shared projects can support relationships. In some situations, relationships can become, at least for a while, projects.

Nevertheless, we talk about relationships and projects in different ways, using different vocabularies. For example, we talk about many projects in goal-oriented terms. You're making *progress* on the violin; you *failed to meet* your mileage target on your bicycle. Such goal-oriented language is a poor fit for most relationships.

I don't intend my treatment of relationships and projects as separate categories to be an implicit claim that a unified account is impossible. Rather, because it is *easier and clearer* to speak of projects and relationships as different things, I will.

In *A Treatise of Human Nature* it is friendship and sociability that Hume credits with pulling him back from the brink of skeptical crisis³ :

Where am I, or what? From what causes do I derive my existence, and to what condition shall I return? ... I am confounded with all these questions, and begin to fancy myself in the most deplorable condition imaginable....

Most fortunately it happens, that since reason is incapable of dispelling these clouds, nature herself suffices to that purpose, and cures me of this philosophical melancholy and delirium.... I dine, I play a game of backgammon, I converse, and am merry with my friends; and when after three or four hours' amusement, I would return to these speculations, they appear so cold, and strained, and ridiculous, that I cannot find in my heart to enter into them any farther (T 1.4.7.8; SBN 268).

More recently, social psychologists have taken an interest in happy people, and have found plenty of evidence to support the armchair view that close relationships like friendship are crucial to our happiness.

Michael Argyle, one of the pioneers of psychological research on happiness, writes that “social relationships have a powerful effect on happiness and other aspects of well-being, and are perhaps its greatest single cause” (Argyle 2001, 71). Ed Diener found the same thing.

³ Throughout this dissertation, quotes from Hume’s *A Treatise of Human Nature* and *An Enquiry Concerning the Principles of Morals* will be cited in shorted form. Citations from the *Treatise* will be prefaced with a ‘T’ and citations from the *Enquiry* with “EPM.” These prefixes are followed by section and paragraph numbers, and page numbers in the Selby-Bigge/Nidditch (SBN) edition of Hume’s works.

Diener and Seligman (2002) found in their study of very happy people that every single one of them had excellent social relationships. Quantity and, more importantly, quality of friendships correlate positively with happiness, and perceived loneliness is robustly linked to depression. In light of this and other parallel findings, Reis and Gable (2003) have suggested that good social relationships may be the single most important source of happiness. It must be true that “it is man, who is essential to man’s happiness” (Tatarkiewicz, 1976, p.130), and as much as some may believe that hell is other people, so, apparently, is heaven (Kesebir and Diener 2008, 122).

Argyle agrees that it is the *closeness* of the relationship that is key:

Close friends are a particular source of happiness. Weiss (1973) found that to avoid loneliness people needed a single close relationship and also a network of relationships. To form a close relationship involves an increasing level of self-disclosure, and without it people will still be lonely (Argyle 2001, 75).

And Ed Diener, summarizing one result of his studies of very happy people:

Very happy people have rich and satisfying social relationships and spend little time alone relative to average people. In contrast, unhappy people have social relationships that are significantly worse than average. One might conjecture that good social relationships are, like food and thermoregulation, universally important to human mood (Diener and Seligman 2002, 83).

The upshot of the recent psychological research is corroboration of the views of Aristotle, Cicero, Montaigne, Hume, and many others: at least one close relationship seems a necessary condition of a life's going well, and in relationships *consists* a significant aspect of a life well lived.

2.2 Personal projects

By “personal projects” I mean concrete activities like gardening, learning the violin, advancing the space program, bringing products to market, holding political office, helping struggling school kids, restoring wetlands, and so on. Personal projects are large-scale endeavors that unfold over time, and around which we organize our lives. Doing the dishes is not a personal project (though keeping a tidy house might be). Going for a run around the lake is not a personal project (though attaining or maintaining physical fitness might be).

Bernard Williams' ground projects—those identity-forming projects without which we might not find life worth living—are examples of personal projects (Williams 1981). But there are many personal projects, as I intend the term, that, despite being meaningful and satisfying, do not rise to the level of Williamian ground projects. Aspects of our circumstances beyond our control can change what personal projects are available to us, and usually we can adapt. An injury can leave us unable to enjoy projects that require easy mobility. A job transfer can end our engagement with projects that require organizations or facilities located in the place we are leaving. In these situations, we can often adopt different projects, and this process need not be psychologically devastating in the way losing Williams-style ground projects would be. The personal projects we had, both before and after adapting to our changing circumstances, can make an important

contribution to our well-being, despite not being as crucially important to our identity as ground projects.

Susan Wolf describes meaningful lives as “lives of active engagement in projects of worth” (Wolf 1997, 209). This notion of active engagement in projects of worth is roughly what I intend by personal projects. She develops the notion like so:

A person is actively engaged by something if she is gripped, excited, involved by it. Most obviously, we are actively engaged by the things and people about which and whom we are passionate. Opposites of active engagement are boredom and alienation. To be actively engaged in something is not always pleasant in the ordinary sense of the word. Activities in which people are actively engaged frequently involve stress, danger, exertion, or sorrow (consider, for example: writing a book, climbing a mountain, training for a marathon, caring for an ailing friend). However, there is something good about the feeling of engagement: one feels (typically without thinking about it) especially alive (Wolf 1997, 209).

The basic idea is that there are some projects that are potentially engaging or satisfying in the way Wolf describes, and some that are not. Grass-blade counting has no potential to be psychologically satisfying for most people. The same is probably true of passive TV viewing or video-game playing, morphine, sudoku, etc. But the range of possible engagements that are potentially satisfying to *someone* is vastly larger than the number of engagements that would be satisfying for any particular person. An important part of life is figuring out which of these many potential engagements are the ones that are especially important to us. Wolf, again, casts this in terms of meaning.

The idea is that in a world in which some things are more worthwhile than others, meaning arises when a subject discovers or develops an affinity for one or typically several of the more worthwhile things and has and makes use of the opportunity to engage with it or them in a positive way (Wolf 1997, 211).

Similarly for personal projects. Personal projects arise when an activity capable of supporting active engagement over time sparks the interest of a particular person. They are the activities we *care* about, the activities that are rewarding, or satisfying, or, in Wolf’s terms, meaningful.

The empirical evidence is accumulating that Wolf is right, and that engaging with the world through personal projects is centrally important to human happiness. In an overview of the current state of research, Kesebir and Diener write that “decades of research reveals... that happiness primarily emanates not from the ceaseless pursuit of pleasure, but from striving for and making progress towards goals derived from one’s most-prized values” (Kesebir and Diener 2008, 121).

Another line of research that reveals the importance of engagement with personal projects is Mihály Csíkszentmihályi’s work on *flow*. Flow is a state of intense absorption with an activity—the kind of absorption in which we lose track of time, and lose track of ourselves.

Being “in flow” is the way that some interviewees described the subjective experience of engaging just-manageable challenges by tackling a series of goals, continuously processing feedback about progress, and adjusting action based on this feedback. Under these conditions, experience seamlessly unfolds from moment to moment... (Nakamura and Csikszentmihalyi 2002, 90).

Being in flow is a great feeling—unbeatable, by Csikszentmihályi’s lights:

[T]he best moments in our lives, are not the passive, receptive, relaxing times—although such experiences can also be enjoyable, if we have worked hard to attain them. The best moments usually occur when a person’s body or mind is stretched to its limits in a voluntary effort to accomplish something difficult and worthwhile (Csikszentmihalyi 1990, 3).

Achieving flow requires a balance between the ability level of the actor and the challenge level of the activity. “The balance is intrinsically fragile. If challenges begin to exceed skills, one first becomes vigilant and then anxious; if skills begin to exceed challenges, one first relaxes and then becomes bored” (Nakamura and Csikszentmihalyi 2002, 90).

I have characterized personal projects as endeavors that unfold over time, and around which we structure our lives. Personal projects like learning an instrument, tending a garden, playing chess, or playing a sport are indispensable for finding the balance between challenge and ability that fosters flow.⁴ This is because the projects themselves are variably challenging. If we become skilled enough at the violin that the songs in the beginner books bore us, we can move on to the intermediate books and have a new set of flow-fostering challenges.

The experience of flow is both a mark of a life that is going well, and the result of engagement with personal projects that are going well. And so I take Csikszentmihályi’s research to be further empirical support for the armchair view that success in personal projects that are of special interest to us is an important part of well-being.

⁴ Indeed, Csikszentmihályi’s research has largely focused on people who are engaged in personal projects: rock-climbers and other athletes, dancers and visual artists, surgeons, chess-players, and so on (Nakamura and Csikszentmihalyi 2002, 89).

2.3 Other candidates

I'm open to the possibility that personal relationships and projects are not the only founts from which well-being springs. Psychologists certainly discuss others, including health, wealth, religion, and mood. However, even the strongest candidates for additional, irreducible sources of well-being—health and mood—are likely to be much less important, relative to relationships and personal projects, for theorizing about the virtues.

2.3.1 Wealth

With respect to wealth, the apparent implications of the research are not too surprising. Although “increased income contributes significantly to happiness at low levels of development across nations, once the threshold of around US\$10,000 annual per capita income has been passed, there is not a strong correlation between wealth and life satisfaction” (Kesebir and Diener 2008, 122). The obvious way to read this result is that a certain amount of financial security is necessary for self-determination. Once that threshold is crossed, and we have enough wealth to allow us the freedom to make and carry out choices, then wealth makes no distinctive contribution to our happiness. What makes us happy is pursuing our values—including projects and relationships—and to secure happiness we need enough money to do that. We don't need more. So wealth is not a good candidate for an additional, distinct source of well-being.

2.3.2 Religiosity

Some social psychological research indicates that in some countries (the United States among them) actively religious people are happier than their non-religious

peers (Poloma and Pendleton 1990). If there is indeed a causal link here, it seems possible that any contribution religion makes to well-being in these countries is via the provision of a venue in which it is easier to pursue projects and find compatible friends. After all, churchgoers have access to a larger group of possible friends than many of their more isolated, non-churchgoing peers, and congregations provide a framework of support for the projects of their members—church musicals, softball leagues, poverty relief via soup kitchens, and so on—and this kind of social support is often less available to non-churchgoers.⁵

I see no reason to suppose that these sorts of explanations can't fully account for the higher average subjective well-being of religious people. More, given the common phenomenon of happy atheists and non-churchgoing agnostics, it can't be the case that religious engagement is *necessary* for well-being in the same way projects and relationships are. (It's easy to be happy while being totally isolated from religion. It's difficult or impossible to be happy while being totally isolated from human relationships.) Thus I don't consider religious engagement an additional, distinct source of well-being.

2.3.3 Health

A more likely candidate for a third item on the list is *health*. The case for health as partially constitutive of well-being is obvious and strong: having the flu is a drag. It's much better to be healthy. And though part of the reason the flu is a drag has to do with the way the sickness affects projects and relationships, these effects don't fully account for the flu's badness. Even if I wasn't planning to see anyone or do anything that day, I'm still worse off if I have the flu.

⁵ Ferriss 2002 offers explanations along these lines, as well as others, in interpreting higher self-reported levels of satisfaction among the actively religious.

Nevertheless, I hesitate to add health to the list of sources of well-being. The problem is that talk of ill-health is unavoidably intertwined with talk of disease and disability. Over the last decade or two, the disability-rights community has made real strides in revealing problematic aspects of many commonly held views about the connection between disease, disability and well-being.⁶ Representative of the best of this work is Harriet McBryde Johnson, a disability-rights attorney who was dependent on a wheelchair from childhood, due to a congenital neuromuscular disease. She makes the case that health is not a contributor to well-being, at least in the ways we typically assume it is.

Johnson recounts an email exchange with Peter Singer, who has claimed parents are morally justified in killing disabled babies in situations in which they can replace them with non-disabled babies. Johnson offers Singer an analogy: is it morally permissible to kill mixed-race babies for the same reasons Singer thinks it is permissible to kill disabled babies? Singer responds that, no, it would be horrible to kill mixed-race babies. “What’s the difference?” Johnson asks.

She paraphrases Singer’s response. “Preferences based on race are unreasonable. Preferences based on ability are not.”

It is Singer’s belief that what justifies his judgment that a preference for white babies is unreasonable is that mixed-race babies are in no inherent sense “worse off” than white babies. While there may be challenges faced by mixed-race children that white children do not face, those challenges all have their origin in social factors like bigotry, not in factors internal to the child. On the other hand, Singer says, his judgment that a preference for non-disabled babies is reasonable because a disabled person is inherently worse off than a non-disabled person. Johnson responds:

⁶ For sustained treatments of health, disability, and well-being, see Wendell 1996, *The Rejected Body* and the first third of Nussbaum 2006, *Frontiers of Justice*.

Are we “worse off”? I don’t think so. Not in any meaningful sense. There are too many variables. For those of us with congenital conditions, disability shapes all we are. Those disabled later in life adapt. We take constraints that no one would choose and build rich and satisfying lives within them. We enjoy pleasures other people enjoy, and pleasures peculiarly our own. We have something the world needs (Johnson 2003).

We all have limitations. Some of us are tone-deaf and others color-blind. Some are bad dancers and others dyslexic. Some need a powered wheelchair to get around and others are terrible cooks. We are all capable of building good lives within our confines. Johnson’s point is that when the confines of the powered wheelchair are singled out by Singer (and many others) as especially unbearable confines, this is a form of chauvinism. It certainly doesn’t reflect the experience of people like Johnson, who build lives within the confines of a wheelchair just as well as able-bodied people build lives within the confines of legs.

Whether or not you find her convincing (I do), Harriet McBryde Johnson succeeds, at least, in revealing the claim that health is necessary for well-being is more *controversial* than it is often taken to be. This is why I prefer to leave health off of the list of sources of well-being, if possible. The gray areas and controversies that follow on any discussion of health, ability, and well-being make a tactical dodge appealing. As luck would have it, a tactical dodge is available.

For the most part, the aspects of our health we have some control over—our health-related *practices* like exercising and eating well—are themselves *projects*. Given that most of us have (or ought to have, for our own sakes) an interest in maintaining practices conducive to our physical health, the importance of health is substantially preserved, even if it isn’t on the list of fundamental sources of

well-being. For most of us, the traits of character that promote health will be among our egoistic virtues because securing or maintaining health is, or ought to be, for most of us, among our personal projects.

The point of a “practical” account of well-being is to find sources of well-being that will be uncontroversial on any plausible theory of well-being. The disability-rights community has highlighted enough problems with our folk understanding of health that its status as a *source* of well-being is appropriately seen as controversial. So I’ll leave health aside, acknowledging the possibility that, if it turns out that health should be on the list of fundamental sources of well-being, it might require some changes to my account of the virtues. Because many character-related aspects of health-as-source-of-well-being are captured by health-as-personal-project, I hope these changes, if needed, will be small and at the margins of my account of the virtues.

2.3.4 Mood

Another likely candidate for a distinct contributor to well-being is mood. Irritability, depression, and other negative moods can tarnish the sheen of successful projects and flourishing friendships, just as confidence and good cheer can enhance it. Though mood often correlates with how well our projects and relationships are going (interacting with friends usually puts us in a good mood, e.g.) it doesn’t always. We can find ourselves prone to melancholy, or annoyed by pet peeves, even when things are going well for us. To the extent that mood can be disconnected from the factors that usually affect it, mood seems like a separate aspect of well-being.

In this case, my reason for keeping mood off the list is purely tactical: there are smart people who view mood as irrelevant to well-being. (Consider notoriously

melancholy Wittgenstein’s reported last words: “Tell them I’ve had a wonderful life” (Monk 1990, 579.) Though I think mood skeptics are wrong—I think mood matters—I leave it off the list in order to avoid needlessly contradicting people like Wittgenstein.

And contracting melancholy Wittgenstein *would* be needless. The research I’ve already cited shows a strong correlation between success in relationships and projects on the one hand, and self-reports of happiness on the other. This isn’t surprising; when things are going well for us we tend to feel good. Wittgenstein’s wracked-but-good life is an outlier—this is part of what makes his last words so notable. The cases in which a good life and a good mood *systematically* diverge will be rare and strange.

Moreover, much as with health, I believe that if it becomes clear that mood *should* be part of my practical account of well-being, the changes required to add it will be small and at the margins. Consider a trait like cheerfulness. If mood counts as a source of well-being, it’s easy to class cheerfulness a virtue. But even if we only count relationships and projects as sources, it’s easy to see something good in cheerfulness: it’s a pro-social trait, and we’re social creatures. In short, focusing on relationships and projects exclusively still captures some aspects of what’s good about a good mood.

And so for now, in order to leave the practical account of well-being as uncontroversial as possible, I leave mood out.

2.4 Projects, relationships, and theories of well-being

The claim that well-being consists, practically speaking, in projects and relationships, is compatible with a broad range of theoretical accounts of well-being. That is, adherents to most of the plausible theories of well-being could endorse the practical recommendation that, to secure (whatever their definition of) well-being, we should seek to do well in relationships and projects that are especially important to us.

Put another way, consider the sentence: “projects and relationships are crucial to a person’s well-being because....” Different theories of well-being might fill in the blank at the end of the sentence in different ways. For a hedonist, projects and relationships are important because they bring pleasure. For an objective list theorist, projects and relationships are important because they are fundamental human functionings. But no matter how a given theory of well-being explains the *reason* relationships and projects are important, any plausible theory of well-being will accept the first clause of the sentence: projects and relationships are crucial to well-being. Let us quickly run through some leading theories of well-being and convince ourselves that they do, indeed, acknowledge the importance of projects and relationships.

In her overview of philosophical and psychological research in well-being, Valerie Tiberius identifies four basic philosophical approaches to the definition of well-being, all of which have some degree of analog in psychology: hedonism, preference-satisfaction theories, objective list theories, and life-satisfaction theories (Tiberius 2006, 494).

Hedonistic theories of well-being are no longer popular among philosophers.

Pleasure alone seems far too thin a concept to capture what we mean when we talk about lives well lived. While simple pleasure may be an aspect of most or all well-lived lives, pleasure alone can't capture things like achievement, meaning, and satisfaction. These are things that we want out of life, even though they are sometimes unpleasant, in the crude sense.

Nevertheless, it's clear that we usually do derive pleasure from our projects and relationships. Pleasure, in fact, is a major motivator urging us to engage these things in the first place. We seek out friends because they are fun to have. We take up hobbies because they bring us pleasure. A hedonist, then, would recognize projects and relationships as important features of a pleasant life.

The explanatory thinness of hedonism led to a more recent generation of theories of well-being. Projects and relationships sit comfortably at the center of this later generation of theories, too.

Preference (or desire) satisfaction theories are the direct offspring of hedonistic theories. The problem with hedonism is that it cannot account for the things we want for ourselves *other* than pleasure. Preference satisfaction theorists define well-being as having our preferences satisfied, whatever those preferences may be. Surely we often desire pleasure, but we might be willing to forgo pleasure to secure financial stability for our families, or lasting fame for ourselves. Preference satisfaction theories hold that well-being is a matter of satisfying our preferences, whether those preferences are for pleasure or something else.

Given that, as a simple matter of fact, most of our preferences will have to do with our relationships and our projects—after all, these are the things we care about the most—to do well in relationships and projects will be tantamount to satisfying our preferences. Thus I suppose the preference theorist can subscribe to the practical advice: to do well in your life, do well in your relationships and

projects.

Objective list theories maintain that well-being is a matter of achieving certain kinds of distinctively human functioning. Tiberius identifies the leading contenders in psychology and philosophy:

Two such programs [in psychology] are the Self Determination Theory of well-being, which posits three basic human needs for autonomy, competence, and relatedness (Ryan and Deci), and Carol Ryff's multidimensional account, which posits six basic aspects of human actualization: autonomy, personal growth, self-acceptance, life purpose, mastery, and positive relatedness (Ryff and Singer).

These accounts overlap significantly with the front runner in objective philosophical accounts, Martha Nussbaum's capabilities approach. Nussbaum postulates ten human functionings that are essential to living well: life; bodily health; bodily integrity; senses, imagination, and thought; emotions; practical reason; affiliation; other species; play; and political and material control over one's environment (Tiberius 2006, 495).

The core importance of relationships and projects is clear in these objective lists. Many of the items on Ryff's and Nussbaum's lists have to do directly with particular kinds of projects and relationships: personal growth, mastery, positive relatedness, affiliation, play. Other items on these lists are arguably prerequisites for successful relationships and projects: autonomy, bodily integrity, control over one's environment, etc. Thus can objective list theorists subscribe to the practical advice: to do well in your life, do well in the relationships and projects that are specially important to you.

Life-satisfaction theorists (Sumner and Tiberius on the philosophy side, Diener on the psychology side) understand well-being as the retrospective endorsement of the circumstances of your life. To be satisfied with how my life has gone requires both my judgement that I've been pursuing the right values, and my judgement that I've done well by those values. From both the armchair and the psychological literature, it seems that the things valuable to us are broadly classifiable into the categories of relationships and projects. Success in the relationships and projects that are of special importance to me is exactly the condition that will yield the retrospective endorsement that life-satisfaction theorists define as well-being.

Here, then, is where we stand. Any *theory* of well-being can sign on to the *practical* recommendation: in order to do well, figure out what projects and relationships *really* make you happy, then do those things well. Projects and relationships thus provide us with a serviceable proxy for well-being. This proxy allows us to make progress in discussing the relationship between character and well-being, even in the absence of a widely accepted theory of well-being.

2.5 Character egoism

What is the nature of the connection between well-being and the virtues of character?

It cannot be the case that the virtues *constitute* or *are sufficient for* well-being. The existence of terrible injustice and terrible luck establish the insufficiency of character for well-being. Imagine a wonderful man—whatever your definition of wonderful—who is kidnapped, enslaved, and regularly tortured until he finally dies. A life of slavery and torture seems clearly enough *not* to be an example of a life well lived, no matter the quality of character that suffers it.

For the same sort of reason, it can't be the case that the virtues are *necessary* for well-being, either. Vices don't *guarantee* misery any more than virtues guarantee well-being. These characters—the thriving villain and the miserable saint—are rare. Nevertheless, their possibility is enough to establish that the virtues are neither necessary nor sufficient for well-being.

Instead, I suggest the virtues *reliably promote* well-being in the sense Rosalind Hursthouse outlines in Chapter 8 of *On Virtue Ethics*: cultivating the virtues is the “best bet” for attaining a good life, much like adopting a healthy lifestyle is the best bet for being healthy.

Consider physical health. Though some degree of good luck is required for me to be physically healthy (I need to be lucky enough to avoid lightning strikes, for instance), adopting a healthy lifestyle increases the chances that bad luck will affect me less and good luck more. Though I might be predisposed to develop cancer, I can increase the odds that I will develop cancer by taking up smoking and I can decrease those odds by giving up red meat.

A healthy lifestyle is a collection of healthy practices, and we can define healthy practices like so: they are practices that reliably benefit one organ or system without inflicting comparable or greater harm on another. Jogging, though it sometimes damages the knees, greatly benefits the heart. It is a healthy practice. Smoking, though it helps prevent weight gain, does great damage to the lungs. It is an unhealthy practice.

I believe the relationship between virtue and well-being is similar. If I am to live a happy life, some amount of good luck is required. I must be born into a decently functional society—a society that supports the possibility of rewarding relationships and projects. If I am born a slave in a cruel society, my prospects for well-being are curtailed or eliminated. But, as in the case of physical health, there

are factors internal to me that interact with external factors. If I am born into circumstances that present me with opportunities for friendship, I can increase the chances that I will have friends by being friendly. And I can minimize those same chances by being mean-spirited.

Thus, I propose the following as the kernel of a theory of the virtues:

The virtues for a person are those traits of character that reliably promote success in the relationships and projects that are (or should be, for her own sake) important to her.

Note that the word “success” could potentially be misleading if it connotes “winning,” or any other objective success conditions. Few projects and fewer relationships are structured around objective goals, and so don’t have success conditions at all. By “success” I mean “doing well by the standards of whatever value you are pursuing.” Success in keeping a flower garden might consist in a collection of vibrant and unusual blossoms. Success in a marriage might consist in a stable relation of mutual care and respect.

Note also that this “practical account” of well-being, which focuses on relationships and projects, only applies to normal human beings. I mean “normal” to be broadly inclusive, here, embracing the overwhelming majority of human beings. However, projects and relationships simply won’t work as a proxy for well-being in strange outlier cases like human/lion hybrids, extra-terrestrials, and (perhaps) psychopaths. Such outliers are likely to be better served by a different practical approach to well-being—one that features considerations other than relationships and projects.

Note also that the full practical advice for doing well is really two-part: 1) Find out what kinds of projects and relationships *really can* make you happy. 2) Do well in those things. I will leave the first of those two items totally untouched.

There are many ways we might err in our judgments about what is important to us. We lack perfect epistemic access to ourselves. We tend, when ranking our interests, to overrate experiences we've had and under-rank those we haven't. We are imperfect judges of character, and so sometimes pursue relationships with people who are a poor match for us. We are social creatures, and may respond to social pressures to seek projects and relationships that aren't the ones we're suited for.

Although virtues of character will undoubtedly be important to the process of figuring out what projects and relationships *really are* the ones that matter to us, I won't say much about it. Figuring out how best to get to know oneself is a project for a different dissertation. The parenthetical clause in my characterization of the virtues—we're concerned with traits that promote projects and relationships that are (*or should be, for our own sakes*) important to us—is intended to acknowledge and avoid the difficult self-knowledge task set before us by item one.

The project I take up in the following chapters is this. Suppose that a normal human being has discovered what projects and relationships really can make her happy. Now, what traits of character reliably promote success in those relationships and projects, whatever they may be? My claim is that the answer to that question looks much like the traditional slate of virtues. If you want to have the best chance at doing well in the projects and relationships that matter to you, you should be kind, courageous, honest, generous and so on.

This is to suggest that the virtues *just are* the traits of character that promote the well-being of their possessor; this is to suggest an egoist theory of the virtues. Before character egoism can get off the ground, we must put to rest the widespread worry that no variety of egoism could be a satisfactory theory of ethics. Putting that worry to rest is the subject of the next chapter.

Chapter 3

Ethical Egoism and Character

Egoism

Making a man happy is quite different from making him good and making him prudent or astute in seeking his own advantage quite different from making him virtuous.

Groundwork of the Metaphysic of Morals

KANT

According to ethical egoism, what makes an action right is its tendency to benefit the actor; an action is wrong to the extent it forgoes benefits, or otherwise sacrifices the interests of the actor. That is, ethical egoism holds that the morally right action, in any situation, is the action that reliably promotes the actor's long-term self-interest.¹ Few moral theories are as enthusiastically scorned. Confidence in

¹ Ethical egoism is usually presented as an optimific theory—that is, right actions are all and only those actions that *maximize* the actor's long-term self-interest. But satisficing forms of ethical egoism are certainly available. The advantages of the “reliably promotes” formulation are 1) it remains neutral between maximizing and satisficing forms of ethical egoism and 2) it maintains the parallel with my preferred formulation of character egoism.

egoism's falseness runs so high that it isn't unusual to find philosophers arguing against other moral theories by tying them to ethical egoism. If theory X can be shown to reduce to egoism, or depend on egoism, or be rooted in egoism, this is enough to suggest that theory X inherits egoism's hopelessness.

Thomas Hurka ran this tactic against eudaimonistic (or flourishing-based) virtue theories in *Virtue, Vice, and Value*. All such theories hold

that a person has reason to act rightly only or ultimately because doing so will contribute to her own flourishing. If she believes this theory and is motivated by its claims about the source of her reasons, her primary impetus for acting rightly will be a desire for her own flourishing. But this egoistic motivation is inconsistent with genuine virtue, which is not focused primarily in the self.... Someone motivated by the theory's claims about reasons will therefore be motivated not virtuously but in an unattractively self-indulgent way (Hurka 2001, 246).

Hurka's view, then, is that eudaimonistic virtue theories are nothing but egoism applied at the level of character and therefore must be wrong.

Julia Annas, in turn, argues that Hurka misunderstands the Aristotelian approach: Aristotelian virtue ethics is *not* fundamentally egoistic. In so arguing, she accepts the conditional premise that "[if] virtue ethics is committed to accepting that it has an egoistic end... [then] it is not a satisfactory ethical theory" (Annas 2008, 209). So a critic and a defender of virtue ethics can agree about this much: egoism is a damning feature of a virtue theory.²

² Other philosophers pursue lines of objection similar to Hurka's. In "The Self-Centeredness Objection to Virtue Ethics" (2006), Christopher Toner catalogs some of them, including Nagel (1986, 195-197), Prichard (1995, 45-46), and Kant (1964, 4:442). Toner's own view is that egoism is a problem for Aristotelian virtue theories, *unless* those theories understand the good of individuals as partially constituted by their communities. In that case, Aristotelian virtue

It is my project, here, to resist the Hurka/Annas consensus. My thesis in brief: ethical egoism is dismissed for *reasons*. Those reasons don't apply to egoism about character; egoism about character traits doesn't inherit the problems associated with egoism about actions. Character egoism shouldn't be dismissed on the grounds that it has "egoism" in its name. Rather, it should stand or fall on its merits.

3.1 Egoism as category mistake

Usually, when we criticize a normative theory, we seek to show that it has implications that conflict with our reflectively endorsed moral intuitions. Kant forbids me to lie to a killer? That's got to be wrong! Utilitarianism requires the punishment of the innocent if that maximizes aggregate utility? That's got to be wrong!³

Instead of being treated in this way, ethical egoism is sometimes dismissed, prior to engagement, as having made a category mistake. This is the sort of consideration that motivates Annas to accept the conditional premise that *if* a theory is egoistic, *then* it is wrong:

The basic idea of ethical egoism is that what ethically justifies what I do, and the way I am, is my own good, where that is distinct from, and potentially in conflict with, the good of others. And we find at once a problem in the idea that this could be an *ethical* position, because of

theories are not *self*-centered, but rather *community*-centered. The relevance to the present work is that Toner, along with all the others, accepts the conditional premise that if a virtue theory is egoistic, then it's wrong.

³ There are other ways to argue against a moral theory than this kind of straightforward intuitive conflict: maybe it is circular, or vague, or self-undermining. All of these critical approaches require engaging the substance of the theory with the goal of showing it fails to account for the moral phenomena.

the very basic thought that ethics is fundamentally about the good of *others*, not my good (Annas 2008, 205).

Annas dismisses ethical egoism by pointing to the familiar distinction between prudence and morality. But the familiar distinction between prudence and morality is exactly the distinction egoists challenge. For Epicurus, Hobbes, Gauthier, and others who could (arguably) be classified as ethical egoists, morality is best understood as a *species* of prudence. Thus, to dismiss ethical egoism in the way Annas does is straightforwardly question-begging.

It's clear that ethical egoists take themselves to be offering a theory of the moral data; they are concerned with accounting for *when* and *why* we have a moral obligations to keep promises, refrain from violence, etc. The ethical egoist answer is that we have these obligations *because* these are the behaviors that are in our long-term self-interest. Thus ethical egoists join a long list of core concepts purported to justify the moral data, alongside such entries as aggregate happiness, inviolable principle, divine design, individual perfection, mutual consent, community flourishing, and on and on.

To dismiss any one of these theories as category mistakes, prior to examining the fit between the theory and the moral data, begs the question against that theory. For example, utilitarians suggest that the moral data is best justified by some version of the greatest happiness principle. It would be straightforwardly question-begging for a Kantian to claim that, because morality is "fundamentally about" submitting to inviolable moral rules, and not about evaluating consequences, utilitarianism isn't a moral theory at all. It is incumbent on Kantians to show that utilitarianism's lack of inviolable moral rules has implications that contradict our reflectively endorsed moral intuitions. (And Kantians have engaged utilitarianism in this way, time and again.)

As with utilitarianism, so with egoism. Ethical egoism's critics should show that the ways it goes about deriving duties to others from self-interest (or, more generally, moral imperatives from prudential ones) has implications that contradict our reflectively endorsed moral intuitions. Though it is question-begging to dismiss egoism because it doesn't justify the moral data *in terms of* duties to others, it is perfectly fair to dismiss egoism if it rides roughshod over our reflectively endorsed intuitions *about* duties to others.

Unfortunately for ethical egoism, when we engage it this way we discover that it does, indeed, don the rough shoes.

3.2 Egoism and beneficence

The temptation to dismiss egoism as having made a category error is a reflection of the contortions the theory must endure in order to generate reasons for beneficent actions. If an egoist stops to help a stranded driver, the egoist's moral explanation of *why* she stopped can't develop in terms of the needs of the stranded driver. The ethical egoist's moral explanation of this action, like every other action, must be rooted in her own long-term self-interest. Even when there is an egoistic story to be told (maybe she anticipates a cash reward) these explanations are incompatible with beneficent *motivation*. Even when ethical egoism gets the moral course of action right (you should stop to help the stranded driver) it prescribes the wrong motivation (you should do it for the payoff). Thus ethical egoism fails to account for the moral goodness of beneficence. Ethical egoism fails to justify the moral data-point that *beneficent motives* are sometimes morally required.

Eudaimonistic virtue theories, whether egoistic or not, don't appear to have this problem. Defenders of theories in the Aristotelian tradition usually insist

(rightly, I think) that virtues like kindness, honesty, and justice require genuine concern for others, whether or not these traits contribute to the thriving of their possessor. David Solomon, for example, writes that these other-regarding virtues “both restrict the attention I am allowed to pay to my own wants, needs and desires, and force me to attend to the wants, needs, and desires of others” (Solomon 1988, 434). Even if kindness contributes to my own well-being, kindness entails a concern for others’ well-being, not out of an expectation of future gain, but because that’s just what it means to be kind.

So, ethical egoism’s self-regarding explanation of altruistic duties is laughably wide of the mark. (“My cost/benefit analysis suggests helping you will accrue to my long-term self-interest, and thus have I done this generous thing for you.”) Character egoism has no such problem accounting for other-directed *acts*, because virtues like kindness respond directly to the needs of others. In order for the beneficence criticism of ethical egoism to cross the act/character divide, it must rear its head anew at the level of character. Does it? Are self-regarding explanations of altruistic *traits* laughably wide of the mark?

Suppose Caroline visits you in the hospital and you ask her why.⁴ She says, “because I knew you were sad, and I wanted to do what I can to cheer you up.”

“You are very kind,” you say. “Why are you so kind?”

“I’ve watched people,” Caroline says, “and I’ve tried it both ways. I’m just *happier* being kind—mean people are, in my experience, lonely and unhappy, and so I do what I can to cultivate kindness.”

Caroline’s explanation, in this case, for why she visited you is the right kind of explanation. She visited you because she saw *you* were in need and wanted to help *you*. (This is the David Solomon point, repeated.)

⁴ The visit-to-the-hospital example is adapted from Stocker (1997), to whom I’ll return shortly.

Caroline's explanation of why she values kindness likewise seems fine. That she believes her good character fosters her own well-being can hardly bother you. Her good character is part of what you like about her, and it would be ghoulish to stop valuing it once you learn that it makes *her* happy, too.

Perhaps this is all too quick. Perhaps to be bothered by Caroline's admission that kindness makes her happy, you need be less a ghoul than a philosopher. Here, in the voice of the objector, is how I suppose the worry might develop:

Caroline values kindness *because* she believes it makes her happy. This suggests that if kindness *didn't* make her happy, then she *wouldn't* value it. If she changes her mind about the relation between kindness and happiness she'll work to scrub kindness from her character and the next time you're in the hospital she won't visit you. This is far too unstable an attitude to count as a virtue. What appeared to be kindness is revealed as an impostor—a set of apparently kind behaviors motivated by self-regard and subject to the vicissitudes of Caroline's beliefs about her own interests. Real kindness would be steady. *Real* kindness would persist whether or not it advantages her to have it.

This sort of instability objection either turns on a mistake or it offers up a bullet that's easy to bite. First, note that Bitter Caroline, who believes that kindness is not in her own interest, is either right or wrong in this belief. If she's wrong, and kindness *really is* in her interest, then the instability objection has no force. Bitter Caroline might *believe* kindness harms her, but she's wrong and so by the lights of character egoism still should be kind. If the instability objection assumes that Caroline's beliefs about her own interests are what ground her virtues, whether those beliefs are true or false, then the instability objection rests on a mistaken reading of character egoism.

On the other hand, if the disposition to be kind *really is* systematically contrary to Caroline's interests, then there's something strange about her case. Either her circumstances are very unlike ours, or her psychology is very unlike ours. How else to explain how kindness, in her case, is a systematic liability when kindness, in our own cases, is not? But if Caroline's case is a strange one, we should be careful in evaluating our own attitudes toward it.

Again: *if* Caroline is right that kindness is hostile to her own well-being, *then* her circumstances or psychology are very different from ours. *If* her psychology or circumstances are very different, *then* kindness might not be a virtue for Caroline.

It's the second conditional that is supposed to be a difficult bullet to bite. But it only *appears* difficult because we are generally bad at imagining alien psychologies and circumstances. When we do the hard work of imagining psychologies or circumstances that really are *different*, the instability bullet becomes much easier to bite.⁵

Lucky for us, science fiction authors have logged many hours imagining alien psychologies and circumstances. Take Pierson's Puppeteers, an alien race of Larry Niven's.⁶ Puppeteers are intelligent, sociable aliens, who constitute an important part of the interstellar economy. They have three genders and are non-humanoid: three legs, no head. Their brain is located in a boney dome on their torso. From that hump sprout two flexible stalks—something between human arms and slug antennae—each of which terminates in a single eye and a mouth with rubbery lips, which they use as we use hands.

Puppeteers are, by human standards, huge cowards. They avoid danger at all costs. If they find themselves in a dangerous situation, they flee to the back of

⁵ Sharon Street runs a similar argument at the meta-ethical level (Street 2009).

⁶ Puppeteers appear in many of Niven's novels and short stories, notably in *Ringworld* (1970) and "The Soft Weapon," collected in Niven 1992.

the group. (In fact, the title of the Puppeteers' highest political office is "The Hindmost.") As a defense against accidental injury they avoid edges and corners in their architectural designs—all their constructed surfaces are curved and smooth. If forced to act in a courageous, or danger-facing, way, they are traumatized. Voluntary danger-facing behavior is only observed in otherwise insane Puppeteers—those with severe depression, suicidal or homicidal tendencies, etc.

Now, courage is a virtue for human beings. Is it a virtue for Puppeteers? I, for one, can't imagine claiming so with any kind of confidence. But if the different psychology of Puppeteers means that courage isn't a virtue for them then why, in principle, couldn't we imagine an alien race with an evolutionary history and present psychology that makes kindness no virtue for them? And if we can imagine an alien race for whom kindness isn't a virtue, then why can't we imagine Caroline having that same psychology?

Yes, positing an extra-terrestrial psychology in Caroline's head is... ridiculous. But such a ridiculous story is one way to be sure we're properly imagining a Caroline who is *correct* when she says that being kind undermines her own well-being. And when we look at her case we find that, far from being an absurd implication of character egoism, the claim that dramatic changes in psychology and circumstance could yield dramatic changes in a person's slate of virtues looks to be *true*.

But surely here we can stop and rest, because even if you're unconvinced—even if you think courage is a virtue for Puppeteers and kindness is a virtue for *every* Caroline and instability objections are a serious problem for character egoism—the nature of the conversation suggests we're in new territory. We're debating the merits of character egoism as a moral theory, not stopping and supposing our work is done as soon as we get to the word "egoism." That's what I'm asking for:

engagement with character egoism that treats it as a theory distinct from ethical egoism, a theory that might well have its own problems.

Recapping quickly: ethical egoism can't justify beneficent motives, and thus fails to account for our reflectively endorsed intuitions about duties to others. Character egoism isn't wrong for that reason; character egoism can give an appealing account of beneficent motivations and other-directed duties. This isn't to say that character egoism is problem-free. It's just to say it doesn't have *that* problem. If character egoism inherits a damning feature of ethical egoism, it must be something else.

3.3 Egoism and moral schizophrenia

The motivational troubles of the ethical egoist, discussed in the previous section, are the source of another common criticism of ethical egoism: ethical egoism is self-effacing, and its agents morally schizophrenic.

A moral theory is self-effacing if it maintains that the criteria of right and wrong the theory develops are not, at least sometimes, the considerations that should *motivate* us when we act. More starkly phrased: a moral theory self-effaces when it tells us that, if we are to do the right thing, we *must not* think about right and wrong using the apparatus of the theory (at least sometimes).

In "The Schizophrenia of Modern Ethical Theories," Michael Stocker offers ethical egoism as a clear example of a self-effacing theory. Says Stocker: an ethical egoist will surely recognize that friendship and love are in her long-term self-interest—and therefore morally right by the lights of her theory. But in order to have relationships of friendship and love, "one must care for the beloved and act for that person's sake as a final goal; the beloved, or the beloved's welfare or

interest, must be a final goal of one's concern and action" (Stocker 1997, 69).⁷

One moral obligation, by the lights of ethical egoism, is to find and keep good, close friends *because* such friends promote your long-term self-interest. But if you are to do the right thing, you must not think about right and wrong using the apparatus of the theory. You must not be motivated by the moral obligations the theory hands down, because being thus motivated makes it impossible to achieve the very goal the theory says you out to pursue. As Stocker puts it, for theories like ethical egoism, "to the extent that you live the theory directly, to that extent you will fail to achieve its goods" (73).

Moral schizophrenia offers a possible path out of direct ethical egoism's friend-desert. If you put out of your mind the fact that you accept the truth of ethical egoism, then you can recover the psychological space to be motivated by the interests of your friends.

If we are to be good ethical egoists, if we are to succeed in doing the egoistic thing (by having close and rewarding friendships, for example) we must divide our selves into an egoist part and a friendship-suited part, which is motivated to value people and projects in a non-egoistic way. In order to be successful as ethical egoists, the agents who accept the theory must hide from themselves their acceptance of the theory. Only then can they form the genuine, close friendships that egoism recommends they should form. This divided psychology is what Stocker calls moral schizophrenia.

(Note that not everyone is spooked by the specter of self-effacingness. Sophisticated consequentialists like Peter Railton, for example, are comfortable divorcing their criterion of rightness from moral motivation and decision-making

⁷ Stocker is speaking of *close* friendship, here. We need not care for our (mere) chess or tennis partners as a final goal of our concern. Throughout this section when I speak of friendship I mean *close* friendship—a relationship of mutual or reciprocal *care*.

(Railton 1984). Whether or not self-effacingness is a mark against a moral theory is not a debate we need to settle for the purposes of this chapter. The argument I'm advancing here is that *even if* self-effacingness and moral schizophrenia are damning criticisms of ethical egoism, they are not criticisms of character egoism.)

3.3.1 Moral justifications and caring motives

We believe that relationships and personal projects are *good things*, things we *ought to have*, and we believe it deeply enough that no plausible moral theory could claim otherwise. But most personal relationships and projects worth having come bundled with the idea of *care*. If you don't care about your son for his own sake, you aren't really engaged in a parent/child relationship. You may be a parent in the sense that you're a donor of genetic material and able to take him to R-rated movies, but you aren't engaged in the *personal relationship* of parent because care is built in to the concept of parenthood. If you are to be a parent in substance—not just one place in a two-place legal or biological relation—you must care about your son for his own sake.

In short, the relation of care is non-instrumental. Thus any moral theory that requires its agents to see everything—all people, all things, and thus all projects, all relationships—as instrumentally connected to the moral good will unavoidably undermine its agents ability *to care* about these things.

Any moral theory that required its agents to be *motivated* by its justificatory principle (supposing that justificatory principle doesn't privilege care) will be a non-starter. Caring motives are good, and they are incompatible with the motive of moral obligation. But though few moral theories justify obligations in terms of care, it is also true that few moral theories require their agents to be motivated by the justificatory principle of the theory. Indeed, most theories

divorce moral justification from moral motivation. That is, the agents of most theories can consider the obligations generated by their morality without those considerations crowding out their caring motives. By divorcing motivation from moral justification the agents of most theories can be concerned with doing the right thing, while at the same time caring about other people and things. But there are better and worse ways to accomplish this divorce.

3.3.2 Three ways to divorce justification and motivation

Acrimonious divorce: moral schizophrenia

Stocker has identified one possible way to divorce justification from motivation: moral schizophrenia. In addition to ethical egoism, Stocker offers classical act-utilitarianism as an example of a moral theory that accomplishes the divorce between justification and motivation by means of moral schizophrenia. Act-utilitarianism holds that an action is right if and only if it maximizes the aggregate happiness of everyone affected by it. An act-utilitarian who lives the greatest happiness principle directly will see all people and all things as instruments by which to maximize aggregate happiness. Such an attitude precludes care. If he is to make for himself the psychological space for friendship and other forms of care, the act-utilitarian must hide from himself the fact that he believes aggregate happiness is the only value that justifies morally right action.

Without moral schizophrenia, it isn't possible to be motivated to adhere to direct act-utilitarianism and at the same time to be motivated by care. From the perspective of the theory, people are instruments by which to maximize aggregate happiness. From the perspective of care, people are ends-in-themselves. These two perspectives cannot peacefully coexist in a harmonious or well-integrated psyche.

The only way these perspectives can coexist is if the agents of direct utilitarianism divide their psyches into parts that are unaware of each other: the caring part, and the utility-maximizing part. That's an unappealing outcome, and to the extent that it requires a psychological bifurcation that's difficult or impossible for human beings to maintain, moral schizophrenia is a clear mark against a moral theory. Moral schizophrenia is a bad way to divorce motivation and moral justification.⁸

Amicable divorce: moral boundaries

In order to roll Kantianism into his critique of modern moral theories, Stocker employs a reading of Kantianism that has since become dated. Stocker reads Kantianism as a variety of monistic theory—a theory that sees every situation as an opportunity to discharge a duty. Understood this way, Kantianism would invite moral schizophrenia, because goods like friendship cannot be got if they are pursued for the sake of duty.

A better version of Kantianism, and one I take to be standard today, reads the motive of duty differently. The motive of duty is a commitment not to cross the moral boundaries set out by the Categorical Imperative. So long as we are acting within those boundaries, the theory doesn't insert itself into our values and motives.⁹

For most Kantians today, Stocker's hospital example would unfold like so. Suppose I know you are in the hospital, miserable. I would like to visit you to try to cheer you up, because you are my friend and I value your happiness.

⁸ It isn't clear to me whether or not Stocker himself thinks moral schizophrenia is an inevitable consequence of divorcing justification and motivation. So I'm not sure if I'm offering a clarification or a criticism in what follows.

⁹ For an influential development of this view, see the first chapter of Barbara Herman's *The Practice of Moral Judgment*, "On the Value of Acting from the Motive of Duty".

If you, for some strange reason, were to challenge me to justify, in Kantian terms, the moral permissibility of my visit, I would say something like this: I framed the maxim, “When my friend is in the hospital I will visit her in order to cheer her up.” This maxim passes the test of the Categorical Imperative and so I can be confident that visiting you is morally OK.

In this case, caring motives and Kantian justification can peacefully co-exist. I wanted to act for your sake, tested my desire against my morality, and then acted on my desire.

As with relationships, so with projects. If the categorical imperative procedure is understood as outlining a set of boundaries on permissible maxims, then I have the freedom to care about the things that matter to me. I can tend my garden because 1) I *care* about my plants, and 2) gardening passes the test of the categorical imperative.

Kantianism understood this way—not as a relentlessly duty-obsessed theory, but as a commitment to a set of moral boundaries—is neither self-effacing nor morally schizophrenic. When caring motives are allowed to share the stage with Kantian moral justifications, self-effacingness and moral schizophrenia bow out.

Nevertheless, Kantianism understood this way *does* divorce the values that motivate the agent from the considerations that justify her action as morally permissible. You can’t, after all, be motivated to undertake any particular course of action out of respect for moral boundaries. It is the need of the friend that motivates the visit; it is passing the categorical imperative procedure that justifies it as morally permissible. And so for post-Rawls Kantians, the divorce between an agent’s motivations and her moral justifications is no kind of criticism at all. It is perfectly possible to be motivated by your friend’s interests to visit the hospital, and at the same time to consider whether or not a visit to the hospital passes the

categorical imperative procedure.

We can tell a similar story about some utilitarian alternatives to classical act-utilitarianism. Consider a version of rule-utilitarianism in which contemplation of the greatest happiness principle is reserved for quiet hours. If the rules sketch boundary conditions on acceptable actions, and during quiet hours these boundaries are adjusted to better reflect the greatest happiness principle, then rule-utilitarianism escapes moral schizophrenia and self-effacingness in just the same way post-Rawls Kantians do. Though the agent's motives (care) are divorced from her criterion of rightness (compatibility with the set of utilitarian boundary rules) the divorce has no ill-effects at all. It is perfectly possible to be moved to act by care, and at the same time to check your proposed action against the set of utilitarian rules you endorse.

The upshot: when a moral theory can divorce motivation and justification without self-effacing or requiring moral schizophrenia of its agents, there's no reason to think the mere fact of the divorce is a mark against the theory.

Amicable divorce: virtue ethics

There's another possible utilitarian use of quiet hours. If what I reflect on in my quiet hours—if what I submit to utilitarian scrutiny—is not the set of rules governing my actions, but rather my own values and dispositions, then utilitarianism will not invite moral schizophrenia. If I am moved by your plight and out of a desire to cheer you up I visit you at the hospital, then I have harmony between my values and motives. If, later that week, I pause to think about whether my values and dispositions are consistent with a concern for the greatest happiness, that in no way undermines the fact that when I visited you, I was (non-schizophrenically!) motivated by care. Indeed, periodically reflecting on whether one's life as lived

reflects one's moral convictions is a praiseworthy thing.

An example of this kind of utilitarianism is Julia Driver's utilitarian virtue theory (Driver 2001). Driver does, I think, stay clear of worries about moral schizophrenia. She accomplishes this by moving utilitarian considerations from the level of action to the level of character. In Driver's virtue theory, the lone utilitarian value of maximum aggregate utility is used to generate a catalog of virtues, but the agents who display these virtues are not themselves motivated by a single value—they can value their friends, their art, and their gardens. So, once again, when the ability to care for people and projects comes on scene, moral schizophrenia ducks out.

Although virtue ethics has not historically been thought to be self-effacing, Simon Keller recently leveled this charge against neo-Aristotelian virtue ethics, as represented by Rosalind Hursthouse. Hursthouse's criterion of right action is this: "an action is right [if and only if] it is what a virtuous agent would characteristically (i.e. acting in character) do in the circumstances" (Hursthouse 2002, 28). If we want to know the right thing to do in a given situation, we need only discover what a perfectly virtuous person would do were she in that same situation.

Keller argues that any theory that develops a criterion of right action in this way is self-effacing. That is, just like direct act-utilitarianism, or direct ethical egoism, neo-Aristotelian virtue ethics divorces justification from motivation, and does it in a way that requires agents to be ignorant of the theory's moral justifications if they are to succeed in doing the right thing.

Keller invites us to imagine a situation in which three friends on a camping trip—Arthur, Benjamin, and Christine—all offer shelter to a family of hikers during a storm. The three friends each have a different motivation. "Arthur's

primary motivation is to help out the hikers and relieve their misery.” Meanwhile, “Benjamin’s primary motive is to act generously.” And, finally, Christine “wants to do what the fully virtuous person would do” (Keller 2007, 225-226).

In this scenario, it is only Arthur who is fully generous, only Arthur who does the generous thing as a generous person does it, because it is only Arthur who is primarily moved by the plight of the hikers. “There is something to regret, from the point of view of virtue ethics, about the motives of Benjamin and Christine: their motives preclude them from being fully generous” (Keller 2007, 227). This is the seed of Keller’s version of the self-effacingness critique. Proponents of virtue ethics, he says,

must say that what makes an act right is its being what the fully virtuous person would do, but add that having the governing motive of acting like the fully virtuous person precludes the possibility of being like the fully virtuous person—so it is often undesirable for people to take as their motives the considerations that provide reasons for action.

Virtue ethics, it seems, is self-effacing (Keller 2007, 227).

On Keller’s diagnosis, the source of virtue ethics’ self-effacingness is the divorce between its criterion of rightness, which appeals to virtuous exemplars, and those very virtuous people, who are not motivated by that criterion of rightness when they act. In short Hursthouse’s criterion of rightness cannot motivate right action. In fact, if lived directly, Hursthouse’s criterion of rightness will *inevitably prevent* fully virtuous action.

Keller is right that the exemplar model of right action divorces justification from motivation. On inspection though, we find that what Keller has spotted is an amicable form of divorce. In fact, Keller’s criticism points to the same feature

of virtue ethics that Julia Annas has elsewhere discussed and deemed “perfectly harmless.”¹⁰

Keller’s article identifies a gap between the approach to moral decision-making neo-Aristotelian virtue ethics recommends to Benjamin and Christine, and the actual approach to decision-making we see in the already-virtuous Arthur. The reason it is harmless is that this gap respects a real difference in their moral abilities. “A beginner in virtue will have to try explicitly to become a virtuous person, and to do so by doing virtuous actions; his deliberations will include such thoughts as that so and so is what a virtuous person would do.” A fully virtuous person, meanwhile, “will not explicitly think about, for example, being brave or performing a brave action” (Annas 2008, 212).

Imperfectly virtuous people are not properly sensitive to the moral requirements of some situations they find themselves in. This is what it *means* to be imperfectly virtuous. In the context of moral reasoning, then, direct appeal to exemplars is an heuristic for the imperfectly virtuous, much like training wheels on a bicycle are a crutch for a child. Training wheels would be an impediment to a skilled rider, but they are, of course, intended to be taken off.¹¹

It is true that a fully generous person like Arthur does not appeal to exemplars. But appealing to exemplars does not *preclude* Christine or Benjamin from becoming virtuous. As they develop the skills and habits of generosity, they will

¹⁰ Though Annas discusses what I take to be Keller’s form of self-effacingness, she does it in an epicycle of her response to Hurka’s egoism objection.

¹¹ Annas, in the passage quoted above, illustrates the point with other practical skills.

A skilled plumber or pianist will simply respond to a challenge, without explicit thoughts about good plumbing or playing; the occurrence of such thoughts notably marks the learner and it is a sign of expertise that they are no longer on the scene. Yet the expert can recover such thoughts to convey the expertise to a learner; no problematic split in the self, or in the agent’s practical reasoning, has been introduced.

no longer need to emulate other people. As they develop virtuous habits, they will internalize those patterns of thought and behavior the exemplars model.

Keller has shown that virtue ethics divorces motivation and justification. If this is all it takes to count a theory self-effacing, then Keller has shown that virtue ethics is self-effacing. I think, though, that the lesson to draw from Keller is that self-effacingness alone, when it does not entail moral schizophrenia, is not a mark against a moral theory. Rather than show that virtue ethics has a problem because it is self-effacing, Keller has shown that self-effacingness isn't always a problem.¹²

3.3.3 The upshot

Ethical egoism and character egoism both involve a divorce between justification and motivation, but they accomplish this divorce in different ways. Ethical egoists, if they are to preserve their ability to care, must first divorce their motivations from their moral justifications, and then they must hide from themselves what they believe to be morally justified. Ethical egoists, if they are to have friends, must be morally schizophrenic.

When considerations of self-interest are moved from the level of action to the level of character, moral schizophrenia disappears. The absence of moral schizophrenia appears to be generally true of virtue theories. It doesn't matter if the catalog of virtues is generated with a utilitarian principle, an egoist principle, a perfectionist principle, or something else entirely—agents who accept a virtue

¹² A glib take on the point: imagine the Keller case against training wheels. It is, after all, a fact that only a person without training wheels can ride *as* Lance Armstrong rides. There is something to regret, from the point of view of the Union Cycliste Internationale, about little Billy's training wheels: his training wheels *preclude* him from riding like Lance Armstrong.

This is silly. Although it is true that no one with training wheels rides as Lance Armstrong rides, it is not true that this gives us a reason to regret training wheels.

theory, in expressing their virtues, can value people and projects for their own sakes, and not as instruments with which to promote the moral good.

In short, many people dismiss ethical egoism because they believe, along with Stocker, that accepting it entails moral schizophrenia. Accepting *character* egoism doesn't entail moral schizophrenia. Agents who accept character egoism can care about people and projects without hiding from themselves what they believe to be morally justified.

But this is enough. Even if you're not yet convinced—even if you think Keller has shown that virtue ethics is problematically self-effacing, or that there's no such thing as an amicable divorce between justification and motivation—we're once again having the sort of debate I'm arguing that we *should* have. We're debating the merits of character egoism as a moral theory, not stopping and declaring ourselves satisfied as soon as we get to the word "egoism."

3.4 Conclusion

Ethical egoism is dismissed for reasons. Some of these reasons—such as the claim that ethical egoism makes a category error—are no good. Other reasons are better: ethical egoism mangles its account of duties to others; ethical egoism divorces justification from motivation in a way that leaves its agents at risk of moral schizophrenia. Neither of these criticisms survives the transition from the level of action to the level of character. Neither of these criticisms of ethical egoism is a criticism of character egoism.

I have not run down every serious problem with ethical egoism. That list is long. Thus, this chapter is not a defense of the claim that character egoism is problem-free, or even that it is free of problems that are rooted in egoism. It is

instead an argument that any ghastly inheritance character egoism receives from ethical egoism needs to be demonstrated, not assumed. In the absence of such demonstrations, character egoism should be engaged on its merits.¹³

¹³ Thanks to Valerie Tiberius and Anne Baril for interesting and useful discussions about this material. Thanks to Sean McAleer and Guy Fletcher for exceptionally detailed and insightful comments on earlier drafts.

Chapter 4

Clearing Ground for Egoistic Virtues

The language of egoism is not a natural fit for the language of virtue. Virtue talk is often linked with talk of altruism and self-sacrifice. Talk of egoism is often linked with talk of selfishness and greed. This is at least in part due to the persistent popularity of Ayn Rand, who embraces the term “egoism” in claiming we should radically revise our moral intuitions to bring them in line with the (supposedly deductively provable) goodness of selfishness and greed.¹ Though it develops in some of the same terminology, Rand’s normative project could hardly be more distant from my own. I’m offering an egoistic theory of the familiar virtues, not preaching a new slate of unfamiliar virtues. In my own view, whatever the egoistic virtues may be, we can be confident that selfishness and greed are not among them.

Nevertheless, the fact remains: the fit is poor between the vocabularies of

¹ See, for example, the essays Rand collected as *The Virtue of Selfishness: A New Concept of Egoism* (1964).

virtue and egoism. This poor fit is likely to lead to confusion. In this chapter I want to preempt some anticipated confusions and make clear which prior philosophical commitments are, and which are not, incompatible with character egoism. In the first three sections that follow, I argue that the putative objection to character egoism is in fact a confusion that is best explained away. That is, the objector in each case is objecting to a theory other than the one I'm advocating. In the fourth section, I discuss the ways in which the situationist challenge to character concepts in general is relevant to character egoism in particular.

4.1 “But, Hitler!” and the threat of endorsing vices

4.1.1 Challenge

There are terrible people with terrible projects. Hitler's life, for example, was given over to global military domination and genocide. If a person's virtues are the traits of character that reliably promote success in the projects and relationships that are important to that person, then aren't Hitler's egoistic virtues the traits that allowed him to murder, oppress, and enslave countless people? But it is preposterous to call Hitler's ambition, hate, callousness, and hubris *virtues*. This, already, is a kind of *reductio* of the idea of egoistic virtues.

4.1.2 Response

The first thing to say is that Real Life Hitler didn't live a happy life. There are plenty of reasons why this isn't a controversial thing to say. Hitler *seemed* neither happy nor satisfied. He seemed harried, bitter, and angry. None of us would

want to trade places with Real Life Hitler, not just because he lived an *immoral* life, but also because he lived a *miserable* life. Had Hitler gotten into art school and settled into a life of landscape painting, surely things would have worked out better not just for the world, but for him, too.

That Real Life Hitler lived an unenviable life suggests that, though he was astonishingly successful (by his own lights) in many of his projects, he wasn't pursuing projects capable of securing his own well-being. But if the projects he pursued weren't capable of securing his well-being, then the traits that help him succeed in those projects aren't necessarily virtues. My claim is that we can discover virtues by looking at the traits that promote success in *good* projects. It's harder to draw any conclusions from the traits that promote success in bad projects.

I have more to say on the issue of good and bad projects in the next section. For now, I'll settle for this much: Hitler was wrong about which sorts of projects would secure his own well-being. Genocide didn't make him happy, and so the callousness and cruelty that made him so good at genocide aren't necessarily virtues for him.

Even if I'm right that Real Life Hitler wasn't made happy by his success, this only halfway answers the objection. After all, we can imagine a fantasy version of Hitler who is right when he says that genocide and military domination really are the best projects for him. That is, we can image an Ideally Coherent Hitler. Near as I can tell, ICHitler shares little in common with RLHitler apart from his projects and his name. But ICHitler does press the objection anew: if ICHitler is *right* that these are the best projects for him, then aren't cruelty and callousness the virtues of ICHitler?

We're now faced with a close cousin of Gibbard's Ideally Coherent Caligula

(Gibbard 1999, 145). Gibbard asks us to imagine a version of Caligula who really is made deeply happy by torturing those around him. He's not confused about himself, he's not in the grip of any false beliefs about the world. Does this version of Caligula have the most normative reason to torture those around him?²

Sharon Street argues that ICCaligula does (or, at least, might) have the most normative reason to torture those around him. If this seems hard to swallow, says Street, it's because we haven't taken seriously just how alien, just how other-than-human, ICCaligula would have to be in order to be coherent (Street 2009, 292-293).

I think Street is right, and the lessons she draws from Ideally Coherent Caligula are directly applicable to Ideally Coherent Hitler. Consider what ICHitler would have to be like. If he were brought a deep and peaceable happiness at seeing families torn apart in death camps, of children being gassed... if he drew a calm pride from surveying bodies piled like sacks of rice... if he were made happy by this kind of suffering, he must be wildly different from you and me. At the least, he must lack any capacity for empathy. That is, he must lack a distinctively human part of his psychology. ICHitler, who gasses people with calm satisfaction, is like a lion toying with a wounded child. It doesn't make sense to condemn the lion in moral terms. The lion's psychology is so different from ours that the moral concepts that fit us just don't fit it. Similarly for ICHitler.

Classifying a trait like callousness as virtue-for-ICHitler doesn't entail *praising* it in the way we praise (encourage) socially beneficial traits. We can classify ferocity as a virtue-for-lions and still seek to keep lions out of our communities. We can classify infectiousness as a virtue-for-bacteria and still wash our hands. If ICHitler is right that he's a monster, there's nothing wrong with treating him as

² Gibbard thinks the obvious answer is "no." It doesn't matter how much Caligula likes it, he has the most normative reason to avoid torturing people.

a monster. We can lock him away or exile him, try to condition a change in his behavior, condemn him to dissuade other, similarly alien people from following his course. None of this is incompatible with believing that, as a monster, an unusual set of traits will serve him well. ICHitler's virtues will be monstrous, as the lion's are ferocious and the bacterium's infectious.

The Hitler objection is a close cousin of the Bitter Caroline example I used in the previous chapter, and my response is roughly the same: *if* Hitler is right about his own good, then his psychology is unrecognizably different from the rest of ours. *If* his psychology is unrecognizably different, then his slate of virtues could be unrecognizable, too.

Recall Larry Niven's Puppeteers, the non-humanoid aliens who are, by human standards, terrible cowards. These creatures press the same sorts of questions as ideally coherent human eccentrics. Cowardice is a vice for human beings. Is it a vice for Puppeteers? I can't imagine insisting that it is. But if the different psychology of Puppeteers means that cowardice is no vice for them then, why, in principle, couldn't we imagine an alien race with an evolutionary history and present psychology that makes cruelty and callousness no vice for them? And if we can imagine an alien race for whom cruelty and callousness are not vices, then why can't we imagine one of those aliens visiting Earth, putting on a Hitler suit, and riding a populist wave in 1930s Germany?

Such a ridiculous story—"Alien in a Hitler Suit!"—is one way to be sure we're properly imagining a Hitler who is *correct* in his judgment that genocide and domination are the sorts of things that make him happy. When we do imagine such a story, our pre-theoretic intuitions about the virtues of Alien Hitler (who is an Ideally Coherent Hitler) weaken or crumble.

Thus I'm willing to accept this implication of character egoism: if someone

has a psychology radically different from ours, they might have radically different virtues, too. This can be true even if they otherwise appear to be a normal human.

4.1.3 Discussion

Though I don't think the Hitler objection succeeds as a *reductio* of egoistic virtues, there are some important lessons to be drawn from the case.

First, the Ideally Coherent Hitler case illuminates a class of philosophical commitments that I suspect are incompatible with accepting character egoism. Character egoism is grounded in the well-being of particular people, and the well-being of particular people depends on the values of those same people. Thus, character egoism assumes an attitude-dependent conception of norms. Street summarizes attitude dependence this way:

There are *no* facts about how an agent has most normative reason to live that hold independently of that agent's evaluative attitudes and what follows from within the standpoint constituted by them; instead, an agent's normative reasons are always ultimately a function of that agent's own evaluative attitudes and what is logically or instrumentally entailed by those attitudes in combination with the non-normative facts (Street 2009, 274).

If you side with Gibbard in thinking that ICCaligula and ICHitler *do not* have the most normative reason to pursue the projects that make them happy, then there's little room for virtue egoism to get off the ground. If you are convinced that cowardice is a vice for Pierson's Puppeteers, there's no space for you to entertain the idea of egoistic virtues.

Second, accepting an egoistic theory of the virtues doesn't necessarily entail saying crazy things about real human beings. There might be substantive philosophical disagreement over cases of ideally coherent eccentrics like ICHitler and ICCaligula but, unless you believe Real Life Hitler was ideally coherent, you should expect character egoists will say the same thing every other sane person will say: Hitler had a set of seriously messed up values that made his own life terrible and the lives of those around him much, much worse.

4.2 Bad projects and the threat of circularity

4.2.1 Challenge

Character egoism understands the virtues—that is, the *good* character traits—as the character traits that foster relationships and projects. Clearly, though, there are *good* and *bad* relationships and projects. The previous section concedes as much: genocide is a *bad* project. Fleecing the elderly, running puppy mills, and producing *The Bachelor* are *all* bad projects.

The previous section suggests that the virtues are the traits of character that foster *good* projects. Hitler's cruelty and ambition don't count as virtues because they are traits of character that foster a *bad* project. This looks like circular reasoning. It's surely the case that if Hitler was a bad guy *because* he had bad projects like genocide, and genocide is a bad project *because* it's characteristic of bad guys like Hitler, then we've got a tight and vicious circle.

4.2.2 Response

The threat of circularity is a confusion rooted in equivocal use of “good” and “bad.” In the context of discovering the virtues, “good” and “bad” should be read not as moral judgments, but rather as judgments about a project’s tendency to promote the well-being of the person whose project it is; a good project in this sense is a *good fit for* the person whose project it is.

Bracket the question of immoral projects for a moment, and let’s look at the most common sort of prudentially bad projects.

Typical prudentially bad projects

It should be clear enough that we can make mistakes in judging what is important to us, what we value, what we like. We’ve all made such mistakes, and most of us have made many of them.

Hector Berlioz, the composer best known for *Symphonie fantastique*, was a medical student when he first visited the opera. He describes his first experience of opera as a self-discovery:

I was thus in a fair way to swell the ranks of the medical students, and might have added another name to the long list of bad doctors but for a visit I paid to the Opera. There I saw *Les Danaïdes*, by Saliere. The gorgeous splendour of the spectacle, the rich fullness of the orchestra and the chorus, the wonderful voice and pathetic charm of Madame Branchu... the crashing bacchanal and the voluptuously dreamy dance-music added by Spontini to the score of his countryman, all filled me with excitement and enthusiasm. I was like a lad with the inborn instincts of a sailor, who, never having seen anything but

fishing-boats on a lake, suddenly finds himself transported to a three-decker in mid-ocean (Lebrecht 1985, 119-120).

And so Berlioz dropped out of medical school, enrolled in a conservatory, and proceeded to write music that has stood the test of time.

Mistakes like Berlioz' enrollment in medical school are a universally shared experience in the process of self-discovery. Everyone has tried out a hobby that looked great for a while, but grew boring with time. Everyone has pursued relationships that turned out, in the end, to be mistakes. Of course this happens. Only with perfect knowledge of ourselves and the external world could we hope to avoid this sort of mistake.

Figuring out which projects and relationships *really are* suited to us is one of the central challenges of life. Fortunately, when we realize that one of our projects is a poor fit for us, switching projects is usually relatively easy. If I discover that what I *really* care about is tennis, when all this time I've been playing soccer, switching to the better-for-me project is mainly a matter of trading in my cleats for a racket.³

Most of these discarded hobbies, career changes, relationship breakups, etc, are projects and relationships that are bad in a non-moral, lower-case 'b' sense of "bad." They are bad because they aren't a good fit. They aren't going to promote the well-being of their possessor. Though medical school was a poor fit for Berlioz, it wasn't an *immoral* project in any sense. There's no reason a moral saint couldn't be a doctor, and indeed draw satisfaction from it.

Note that traits that promote success in poor-fit projects and relationships

³ Switching mistaken projects and relationships is usually *relatively* easy. Relative, that is, to the terribly imprudent/immoral species of welfare mistake I discuss next. It can still be difficult or painful to walk away from a project we've invested a lot in. It's *usually* difficult and painful to walk away from a relationship that turns out to be a mistake.

might well promote success in projects and relationships that are a good fit. Thus, poor-fit projects can still be helpful in thinking about virtues. The patience and concentration Berlioz developed studying anatomy probably served him well in studying the scores of the great composers. That is, some of the same character traits he needed to cultivate in order to be a doctor will serve him well in the projects he finds more rewarding. Thus it is that we can make substantial progress in discussing the virtues long before we have a complete psychological analysis of individual happiness.⁴

Immoral projects

We can read virtues and vices off of projects and relationships before we make any *moral* judgments about which projects are praiseworthy and which blameworthy. There is no circularity. But there *are* immoral projects. How can character egoism distinguish immoral projects from projects that are merely a bad fit?

For now, I can only offer the promise of future answers. This is because the virtue theory needs to be in place before we can appeal to it for moral judgments in particular cases. First, the idea of well-being yields an egoistic theory of the virtues, second, the virtue theory yields an account of right action, and finally the account of right action can class projects as permissible or impermissible.

This approach to particular moral judgments is no different than the neo-Aristotelian approach. The difference between the theories is at the top level, in the generation of the catalog of virtues. For both character egoism and Aristotelian virtue ethics, genocide is a bad project because it is a project that expresses vices—a project no decent person would have.⁵

⁴ The same is true for immoral projects. Adapting Kant: coolness may make the scoundrel all the more repugnant, but that coolness is still a virtue (Kant 1964, 62).

⁵ To be clear, in order to talk about immoral projects with clarity and confidence, I first need

However, even before deriving a principle of right action from an account of good character, the egoistic approach gives us *some* resources for distinguishing the usual variety of imprudent projects from projects that are probably immoral.

Correlation between terribly imprudent and immoral projects

There are poor-fit projects that are much, much worse for their possessors' well-being than projects like computer programming or soccer. A few literary examples: Ahab and his revenge on the white whale, Richard York and his ambition for the English crown, Ebenezer Scrooge and his pursuit of profit.

These are all examples of people with projects so ill-suited to their own well-being that the projects actively *harm* the people who pursue them. What makes them qualitatively different from the typical varieties of ill-fitting projects is that they all require, for their own success, the cultivation of dispositions that are detrimental to broad swaths of *other* possible pursuits. In fact, they are detrimental to the very sorts of pursuits that would be more satisfying to the agent if he or she were to undertake them.

I suggest that imprudent projects of this kind—projects that undermine other projects that are better for the person who has them—correlate in many cases with projects that will turn out to be immoral once we've derived a theory of right action from the virtue theory. Let's page quickly through these examples.

Captain Ahab's monomania serves him well in hunting down Moby Dick; it's hard to imagine a well-balanced person pulling off the feat of locating a single

to derive a theory of right action from the egoistic theory of the virtues. This derivation isn't a project I take up in this dissertation. The basic approach of neo-Aristotelian theories is, I think, the right one: what defines right action in a situation is the action that would be undertaken by a fully virtuous person in the same situation. However, adapting this approach for character egoism is non-trivial. Character egoism embraces real moral variety among virtuous characters. Thus, the idea of an exemplar becomes complicated.

whale in the vastness of the sea. But Ahab's monomania confines him to that single project, and undermines every relationship available to him. Ahab's crew understands that if he could abandon his pursuit of revenge, and reacquire the life he had before the whale ate his leg, Ahab would be a happier person. But the project of vengeance has shaped him in a way that makes it impossible for him to adopt the projects that would, indeed, make him happier if he had them.

Richard of York (or, at least, Shakespeare's version of him) cultivates a whole slate of vices in his pursuit of the English crown, and the vices serve him well in that pursuit. He was so far out of the line of succession that, without the heartlessness and treachery that let him murder with impunity, without the dishonesty and injustice that let him lead his friends to ruin, and without the blind ambition that never let him pause for breath, he never would have won the throne. But Richard is neither happy nor satisfied with his success. If he had some of the other of the many projects available to him as Duke of York, he would have been much happier than he turned out to be as King of England. But the very character traits that serve him so well in his pursuit of the crown prevent him from pursuing the sorts of projects and relationships that would have made him happier.

The character of Scrooge is unusual in that we get to see him experience an alternate way of life, and then we get to see his subsequent (totally implausible) conversion. The avarice and hard-heartedness that serve Scrooge so well in his pursuit of profit would make difficult or impossible his transition to the more satisfying roles of friend, benefactor, and teacher. Fortunately for Scrooge, Dickens writes the pernicious dispositions right out of his character. A person as greedy and cold as Scrooge would find it difficult or impossible to adopt the sorts of Tiny-Tim-mentoring projects that would actually make him or her happy.

Like all poor-fit projects, terribly bad projects fail to make their possessors happy. What makes them terrible is that they shape their pursuer's character in ways that make it difficult or impossible to change course, to adopt projects or relationships that *would* lead to happiness.

These terrible projects frequently correlate with immoral projects. It isn't surprising that these terrible projects are also usually immoral, because they typically require the cultivation of *vices*. This is, after all, how they undermine our ability to pursue other projects. Vices just are those traits of character that harm our ability to pursue rewarding relationships and projects.⁶

But, again: terribly bad projects are, first, bad in the sense of "bad for us." This is connected to badness in the moral sense, but the virtues and vices are not defined relative to the moral aspect of these projects. They are defined relative to the well-being aspect.

4.2.3 Discussion

Although there is no circularity in the definition of good projects and good character traits, the preceding discussion focuses attention on an important aspect of the character egoism project that I won't develop in this dissertation: the *structure* of the catalog of egoistic virtues. The complication is easiest to see in the case of vices.

A vice, I suggest, is a trait of character that harms our ability to succeed in the projects and relationships that matter to us (or should, for our own good, matter to us). But this claim can't be as straightforward as it looks.

⁶ The correlation is not always perfect. Ahab's whale-hunting, for example, would not be classed an immoral pursuit by people skeptical of the existence of duties-to-self.

More realistically, it's easy to imagine a teen unsuited for military service enlisting and being permanently harmed by changes of character he must force in order to survive boot camp. His enlistment was not thereby immoral.

Consider the virtues of a playwright and a politician. Some of the virtues required to succeed in one project hamper the pursuit of the other. A playwright needs a level of emotional sensitivity that would be a liability for politicians, who need a thick skin.

Now imagine Amy, who has graduated from college and is trying to decide what career to pursue. She has deeply felt political views, and a career in politics appeals to her. She also has a creative streak and real talent as a writer. Suppose she has to pick one of these careers—being a playwright/politician is not an option.

Suppose Amy chooses to write plays, and cultivates the traits of character that serve her well in that role. She thereby harms her ability to succeed as a politician. This is a life-role she chose to forgo, but nevertheless it's a project that *does* matter to her, a project that *would* make her happy. Thus, Amy's playwright virtues meet the crude definition of "vice" I suggested above: her sensitivity undermines her ability to succeed in a project that matters to her. This is a problem with the crude definition of vice.

The problem is that some kinds of projects are incompatible with each other. The sort of people who excel in one variety of praiseworthy project may be terrible at another variety of praiseworthy project. How, then, can an egoistic approach to the virtues avoid the trap of mistakenly classing competing virtues as vices?

I think the answer is that the egoistic catalog of virtues can't be a flat list. There are some virtues that are more important than other virtues. Some virtues, like kindness, are centrally important to almost every form of rewarding relationship. Kindness is thus a core virtue, and almost everyone will need it if they are to do well in their lives. But not every virtue is so universally beneficial.

Consider a grief counselor with a quick wit and a cheerful disposition. These are traits that could well be desirable for many other people. But if our grief

counselor fails to keep a tight lid on his quick wit and cheerful disposition, we'll judge him to be a bad grief counselor. To the extent that he consistently fails his clients, we might even judge him a bad person.

There are two dimensions along which a vice or virtue can affect our projects and relationships. A vice, for instance, can negatively impact a broad or narrow range of projects, and can do a small or great degree of damage. A trait that is crucially important to a broad range of projects or relationships is a core virtue—one virtually everyone needs to live a good life. A trait that moderately advances a narrow range of projects will only be a virtue for those few to whom those projects are very important.

The traditional (or Aristotelian) virtues seem, by and large, to fall into the former category: they are very important to success in a broad range of projects, and so will be virtues for (most) everyone. A ready wit, on the other hand, is only of value to people who associate with groups of people who *appreciate* a ready wit. A ready wit does nothing for a grief counselor, and may even be a liability. It's a fringe virtue.

Though I think this core/fringe structure is a crucial feature of any plausible catalog of egoistic virtues, I leave its development for future research.

4.3 Context-sensitivity and the threat of instability

4.3.1 Challenge

If the virtues are traits of character that reliably promote good relationships and projects, then what counts as a virtue will change with changing circumstances.

If we move from one set of circumstances to another, it's possible that a different set of character traits would promote success of the same projects. If honesty promotes success in context A, it's a virtue in A. If honesty hinders success in context B, honesty is not a virtue in context B. So if Roderick travels from A to B, his virtues change along the way.

This is a problem for character egoism, because one of our basic pre-theoretic convictions is that the virtues are *stable*. An honest person characteristically tells the truth whether or not he's in a situation in which honesty gets him what he wants.

4.3.2 Response

The stability objection can be developed in two ways. One version of the objection is grounded in a confusion, the other version offers up a bullet I'm willing to bite.

Instability Version One: local changes

One version of the stability worry looks at changes in circumstance that happen moment-by-moment in the life of a person. Context A, say, is Roderick talking with Ligeia, and in this context honesty gets him what he wants. Context B is Roderick talking with Berenice later that day, and in this context honesty prevents him from getting what he wants. Does this mean that according to egoistic virtue ethics, honest dispositions are virtuous for Roderick in the morning, but in the afternoon they're vicious?

No. Character egoism doesn't recommend such plasticity of character, because this isn't how human characters work. For better or worse, our characters *are* remarkably stable over time, whether we want them to be or not. If we take it into our heads that we want to change some aspect of our characters—say, we

get sick of our own impatience and decide to cultivate patience—this is almost always a hard slog. It takes concerted effort over time to change our dispositions, attitudes, default positions. We simply can't be *dispositionally* honest in the morning, and *dispositionally* dishonest in the afternoon.

No theory of the virtues that respects the boundaries of the possible will recommend that human beings change their character from scene to scene in the course of their day. An egoistic virtue theory certainly doesn't. The question character egoism asks is, "what traits serve me well in my life?" not "what traits serve me well at the moment?"

Instability Version Two: global changes

Another version of the stability worry looks at changes across entire worlds. We might develop this objection like so. Consider the virtue of kindness. Now, imagine a world in which acts of kindness disgust everyone who sees them—it actually makes them physically ill. In this world, kindness would make relationships difficult or impossible to form or maintain. So, according to egoistic virtue ethics, kindness would be a vice in this world. But kindness is a virtue. So, *reductio!*

Egoistic virtue ethics would indeed declare kindness a vice in the imaginary disgust-world. I think this kind of instability—which is really just a sensitivity to the way our characters interact with our societies and the external world—is an attractive feature of a virtue theory, and there are plenty of examples of virtue theories that embrace it.

Julia Driver imagines a world in which a malevolent version of Maxwell's demon systematically intervenes to make every attempt at generous action have terrible results (Driver 2001, 79-80). Has Driver's demon made a vice of benevolent dispositions? (Driver, a utilitarian about character traits, thinks the answer is

yes.)

David Hume runs a similar thought experiment. He thinks it would be absurd to maintain the same slate of virtues if you were plucked out of peaceful and orderly Scotland, and dropped in the middle of a violent anarchy.

Suppose, likewise, that it should be a virtuous man's fate to fall into the society of ruffians, remote from the protection of laws and government, what conduct must he embrace? He sees such a desperate rapaciousness prevail, such stupid blindness to future consequences, as must terminate in destruction to the greater number and a total dissolution of society to the rest. He can have no other expedient than to arm himself, to make provision of all means of defense and security. And his particular regard to justice being no longer of use to his own safety or that of others, he must consult the dictates of self-preservation alone, without concern for those who no longer merit his care and attention (EPM 3.9; SBN 187).

Hume, Driver, and I all share the intuition that huge, systematic change in the outcomes associated with our character traits can, and should, change the attitude we take toward those traits. This is easiest to see in fanciful examples like the ones we've just looked at. But I think it is an effect that can sometimes be detected in the real world, too.

One possible example of a virtue that died out as the real world changed is feminine chastity. There was a time when chastity—a dispositional aversion to sex outside marriage—was widely held a virtue. It no longer is. One possibility is that we were wrong then (or now) about the goodness of chastity. But another possibility is that changing circumstances have changed chastity's goodness.

Here, in brief, is the character egoist argument that feminine chastity is a trait that changed its valence with changing circumstances. When Hume classed it as a virtue *of women* it was a trait that could plausibly be connected to women's welfare. The absence of contraception meant pregnancy was a likely consequence of extra-marital sex, and the nature of pregnancy makes it difficult for a mother to escape the burden of child-rearing by denying maternity. ("From this trivial and anatomical observation is derived that vast difference betwixt the education and duties of the two sexes" (T 3.2.12.3; SBN 571).) In short, the potential consequences of extra-marital sex fell more heavily on women, and this asymmetry was reflected in the virtue of chastity.⁷

In the present world, where contraceptives reduce the risk of pregnancy and paternity is almost as difficult to deny as maternity, chastity's status as a virtue is widely questioned, and Hume's claim that it's a *woman's* virtue seems silly or offensive.⁸

If it is true that chastity is a trait that was once, but is no longer, a woman's

⁷ This is a white-washed account of sex-role asymmetries in Hume's day. As Annette Baier makes clear, the sexism of Hume's culture is the main cause of the asymmetry in costs. Hume's argument "depends upon an assumption made but not stated, namely that a society be both patriarchal, so that control of 'expences' is in male hands, and patrilineal, so that property passes through the male line" (Baier 1979, 8). The point remains: social conditions for women were worse in the Eighteenth Century than they are today.

⁸ Todd Calder makes similar comments about chastity in the course of defending an intrinsic account of virtue (Calder 2007, 211). For Calder, chastity was once a form of "loving one's own good." Changes in social circumstances mean chastity is no longer a form of loving one's own good. So chastity is, for Calder, an intrinsic virtue contingent on the details of social arrangements.

Though Calder uses the language of virtue and vice to discuss chastity, his approach seems to me to be a straightforward denial that chastity is a character trait. Instead, Calder is characterizing it as a practice implementing a virtue, and the *practice* changed with the circumstances. (Belching at the table is a practice that implements the virtue of gratitude in some cultures and not in others. No one thinks belching, itself, is a character trait capable of classification as a virtue or a vice.) If we believe that chastity *can* be described as a character trait—a disposition of disinterest in sex, say—then Calder's side-step doesn't weaken it as an example of a virtue that has changed with the times.

virtue, this is good reason to prefer an account of the virtues that is responsive to changes in circumstances. Character egoism is such an account. Sensitivity to circumstances—what a critic might class instability—is a feature of the theory, not a bug.

4.3.3 Discussion

Some discomfort with the idea of the catalog of virtues changing in different circumstances might be due to weak effort when it comes to imagining those different circumstances. After all, the circumstances we find ourselves in don't change much from day to do, or year to year. It *is* absurd to imagine our own slate of virtues changing from morning to afternoon, or even from year to year.

More, cultural variation within our own world rarely seems wide enough to change the sorts of characteristics that help us succeed in projects and relationships. It shouldn't come as a surprise, then, that virtues seem to be pretty stable across cultures in our actual world, too.⁹ Moving from Annapolis to Santa Fe isn't going to make cruelty a virtue. Moving from Annapolis to Moscow won't, either. There just isn't enough variation in human cultures to make major, fundamental revisions to our catalog of virtues.

But a change like Hume's society of ruffians, or Driver's malevolent demon very

⁹ Julia Annas discusses the Virtues Project in her Presidential Address to the Pacific APA:

If you get on the web and go to www.virtuesproject.com, you will find the Virtues Project, an organization which specializes in moral education and conflict resolution, which has been particularly successful in the First Nations areas of western Canada and in Maori areas of New Zealand. It does this by using the language of the virtues, which they have found to be the most effective inter-cultural ethical language. The website features a list of fifty-two virtues which the project has found to be character traits respected in seven world spiritual traditions (Annas 2003).

well might. These drastically different societies—so different they can be difficult to imagine—might drastically change the traits of character that would serve us well in our relationships and projects. So, egoistic virtues might be different in these imaginary societies than they are in our own.

Here, then, is another stop at which some might want to hop off the character egoism bus. Character egoism *is* sensitive to the interaction between characters and the external world, and not everyone is comfortable with considering the external world when theorizing about the virtues. Indeed, the intuition that good people should have generous dispositions, even in the demon world—even when generous dispositions *reliably* lead to bad outcomes—is a key intuition underlying “internalist” virtue theories like Michael Slote’s agent-based approach (Slote 1995).

If you side with Slote in thinking that virtue theories should be concerned exclusively with intrinsic features of traits, and that any appeal to consequences in the external world is a mistake, then there is little room for an egoistic theory of the virtues to get off the ground.

4.4 The situationist challenge to virtue ethics

4.4.1 Challenge

Experimental psychologists have amassed evidence that undermines philosophers’ approach to character.

- Milgram’s obedience experiments show that a little cajoling from people in white coats makes experimental subjects act cruelly. And yes, the white coats matter (Milgram 1974).

- Isen and Levin’s helping experiments show that a windfall dime in a coin-return prompts significantly more generous behavior than no dime in the coin-return (Isen and Levin 1972).
- The “Good Samaritan” experiments at Princeton Theological Seminary show that being early for a meeting makes people sensitive to human suffering, while running late makes people callous (Darley and Batson 1973).

The takeaway from these experiments should be that details of a situation—sometimes even tiny and seemingly irrelevant details—are what yield outcomes. The characters of the people involved make little or no contribution. This view—that situations, not characters, are what yield outcomes—is called “situationism.” If situationism is true, it cuts the legs from below any character-based moral theory, including any egoistic theory of the virtues.

4.4.2 Response

The situationist label obscures the diversity of philosophical positions collected under it. In fact, different strands of philosophical situationism have different implications for virtue ethics. On one end of the spectrum are moderates like Kwame Anthony Appiah, who writes that the situationist hypothesis holds “that, in explaining behavior, we’re inclined to overestimate disposition and underestimate situation. It doesn’t claim that dispositions don’t exist” (Appiah 2008, 50). This variety of situationism doesn’t *threaten* virtue ethics at all, but rather offers boundaries within which virtue ethics must develop if it is to remain consistent with the empirical research of psychologists. Character egoism, which is motivated, in part, by a desire to better capture the interplay between character and circumstances that is part of every person’s well-being, is well-positioned to

develop within these boundaries.

On the other end of the spectrum are situationists like John Doris, whose conclusions purport to strike at the foundations of virtue ethics.¹⁰

John Doris's version of situationism is especially threatening to egoistic virtue ethics. His target in *Lack of Character* is the idea of "robust character traits," which he characterizes like so: "If a person has a robust trait, they can confidently be expected to display trait-relevant behavior across a wide variety of trait-relevant situations, even where some or all of these situations are not optimally conducive to such behavior" (Doris 2002, 18).

The argument of *Lack of Character* is that experiments like those mentioned above show that people *do not* act consistently across trait-relevant situations; indeed, small changes in circumstances can yield large changes in behavior. Thus, *modus tollens*, people don't have *robust* character traits, though they might have what Doris calls *local* character traits—a characteristic way of acting in identical (or maybe very similar) situations.

Virtue theories do assume character traits that are more robust than Doris's local traits. An honest person, for example, can be expected to tell the truth not only to her parents, but also her friends and teachers, not only when she has something to gain by her honesty, but also when she has something to lose. If she only told the truth to selected classes of people, or only when it benefited her, she wouldn't be *honest*. So, unlike Appiah's situationism, Doris's project does threaten most projects in virtue ethics.

We have no shortage of reasons to be skeptical of Doris' conclusions. Sabini and Silver criticize Doris' interpretation of the experimental data. They look closely

¹⁰ Actually, the other end of the spectrum is anchored by Gilbert Harman, who believes that all talk of character should be abandoned (Harman 2000). Most criticisms of Doris (including those I discuss here) apply *a fortiori* to Harman.

at each of the studies Doris cites as evidence, and instead of finding support for the situationist thesis and its broad denial of robust character traits, they find a group of particular discoveries about human psychology: “people’s understandings of the world... are strongly influenced by what they take to be other people’s perceptions of [the] world.” People “are confused and inhibited by the anticipation of embarrassment.” And so on (Sabini and Silver 2005, 559).

Rachana Kamtekar, on the other hand, criticizes the understanding of character Doris depends on in his argument.

The so-called character traits that the situationist experiments test for are independently functioning dispositions to behave in stereotypical ways, dispositions that are isolated from how people reason.... By contrast, the conception of character in virtue ethics is holistic and inclusive of how we reason: it is a person’s character as a whole (rather than isolated character traits), that explains her actions... The virtuous character that virtue ethics holds up as an ideal is one in which [a wide variety of] motivations are organized so that they do not conflict, but support one another. Such an organization would be an achievement of practical reason, and its behavioral manifestation would be cross-situational consistency (in a sense somewhat different from the situationists’) (Kamtekar 2004, 460).

It is true that psychological experiments show that people often act (drastically) differently depending on small features of a situation. Doris infers drastic situation-by-situation variation in *dispositions* as the best explanation of drastic situation-by-situation variation in *actions*. The root problem is that Doris glosses over the obvious difference between the way people act in different circumstances,

and the dispositions that underlie those actions. As a result, his inference to the best explanation isn't compelling.

In the case of Milgram's obedience experiments, for example, we might suppose, ahead of time, that a *kind* person would refuse to turn up an electric current to levels that caused obvious agony. Such an act is cruel, and that should be incompatible with a kind disposition. But when you add an authority figure ordering such an act, you can get loads of people—many of them, presumably, kind—to act cruelly. Doris concludes that people aren't actually kind or unkind, that there's no robust character trait of kindness at all. But this conclusion isn't warranted until we know what it was *like* for the people in the experiment. If it caused them *pain* to act cruelly—if they themselves suffered sympathetically with the subject, or regretted their actions, or felt like they acted out of character—then it seems reasonable to say that they *are* kind people, who are also (excessively) deferential to authority. In this experiment their kindness came into conflict with obedience and lost. And that seems to be exactly what happened in the Milgram experiments. Yes, an appalling number of people obeyed authority; but many of them suffered in the process. These, I should think, were the kind ones. Kind in the robust-character-trait sense.

The upshot is that we can't infer fragmented characters from fragmented actions in the way Doris does. The interplay between various dispositions and circumstances will yield a single observable action, and that single action doesn't always encode all the information about the dispositions underlying it.

4.4.3 Discussion

Doris' view of character as a huge collection of local traits—situation-specific dispositions that don't correlate with neighboring local traits—threatens most

projects in virtue ethics, but is especially threatening to character egoism. If Doris is right, Aristotelians could in principle adapt by greatly expanding their catalog of virtues and indexing traits to situations. But if Doris is right, character egoists will turn out to be incapable of accounting the moral phenomena.

The final chapters of this dissertation are given over to an argument that character egoism can give a plausible account of honesty and justice—two virtues often thought to be fundamentally other-regarding. In both cases, my argument turns on the claim that our dispositions (our character traits) *cannot* be fragmented and compartmentalized in the way Doris thinks they naturally are.

An example that will come up again in more detail in section 7.2: under American racial slavery, white slave-owners treated black people as property, and white people as fellow citizens. This is a clear compartmentalization of *behavior* in neighboring locales. However, I think it is naive to suppose that our psychologies respect the same boundaries as our behaviors. That is, I think it's naive to suppose that cultivating callousness toward the black people who live in the slave quarters doesn't do some collateral damage to your ability to empathize with the white people who live inside your house.

A more mundane example: some people feel that honesty is something owed to people close to you, but not to people outside your circle of familiars. Family members, for example, deserve the presumption of honesty; coworkers don't. I argue that this is an unstable arrangement; the *dispositions* underlying the behavior don't respect the same categories. If you force yourself, over time, to feel *comfortable* lying to co-workers by default, you'll end up feeling more comfortable lying to family members, too. If Doris is right, and compartmentalizing our dispositions is not just easy, but in fact our natural state, these arguments on behalf of honesty and justice collapse, and I see no way for character egoism to make

sense of these virtues.

But again, Doris' entire project is to infer compartmentalized dispositions from apparently compartmentalized behaviors in psychological experiments. This inference is totally without warrant. To the best of my knowledge, though, there is no consensus view among psychologists on the question of how effectively we can compartmentalize our sympathies. I hope that with time, more psychological research will settle this question in my favor. But should Doris be vindicated, the possibility of an egoistic virtue theory that can account for our pre-theoretic intuitions about justice and honesty is obliterated.¹¹

¹¹ Thanks to Michelle Mason and Mike Rohde for discussions about this material.

Chapter 5

A Catalog Grounded in Well-Being

Any moral theory must ultimately be judged by how well it justifies the moral data, our reflectively endorsed intuitions about moral matters, and the clarificatory help it offers when those intuitions are confused. In this chapter I first discuss how to use an egoistic definition of virtue to classify particular character traits as virtues and vices. Second, I sketch some examples that suggest this process can generate a catalog of virtues that saves the moral data. Third, I show that the egoistic approach can contribute to the debate about modesty, a trait whose status as a virtue is controversial.

5.1 Getting a catalog out of the definition

I've proposed this definition of the virtues: they are those traits of character that reliably promote success in the relationships and projects that are (or ought to be, for her own sake) important to their possessor. This definition allows that a virtue

could have some downsides, or that a vice might have beneficial side-effects. This is good, because real character traits interact with the real world this way. People *do* sometimes pay for their kindness; people *do* sometimes profit from treachery.

If we allow our attention to be fixed by a single facet of a character trait, we might be misled in our judgments about that trait. If we look only at the financial success ruthlessness brings an aspiring corporate executive, we might think his ruthlessness a virtue. If we look only at the difficulty kindness gives a teacher in appropriately disciplining his students, we might think his kindness a vice. These judgments would be mistakes. Before we can make a judgment about the moral valence of any given character trait, we must first sort out the tangle of costs and benefits that flow from that trait. Let us return, for a moment, to the analogy of physical health, first discussed in chapter one.

Healthful practices are those practices that reliably promote the health of the practitioner. Given this definition, showing a given practice to be a healthy practice requires two steps: first, we must show that the practice has a positive effect on some organ or system and second, that it doesn't inflict collateral damage on other organs or systems—damage that outweighs the benefit.

An example: an appropriate amount of body fat is partially constitutive of good cardiovascular health. If we can identify a practice that tends to keep our weight within the appropriate range, and that practice doesn't collaterally damage another system, we have identified a healthy practice.

Jogging a few times a week helps keep weight down. Now, jogging does some collateral damage to the body. The worst damage tends to be to the knees; but this damage to the knees is both less certain and less important than the cardiovascular benefits. Thus, because jogging promotes the health of one system, without inflicting comparable damage to another, we can identify it as a healthy

practice.

Smoking also tends to help keep weight down. However, smoking does terrible harm to the lungs. In fact, the negative effects of smoking seem more certain and more serious than its positive effects. Thus, while smoking does indeed promote a trim waistline, it's not a healthy practice. There are many, many ways to control weight that don't come bundled with the risk of emphysema and cancer.

We should approach the classification of virtues in a similar fashion: by looking at the typical impact of a trait on all aspects of a person's life, and investigating whether the sum of effects constitute a net benefit or harm. Thus, if a trait of character tends to promote our ability to enjoy some set of relationships or projects, and that same trait doesn't significantly hamper our ability to enjoy some other relationships or projects, then that trait is a virtue.

There are many different meaningful relationships. A person might be all at the same time a friend, child, partner, and mentor. There are many different sorts of personal projects. That same person might value scientific research, political activism, and the opera. If a trait reliably promotes her ability to enjoy one or more of these sorts of things, without collaterally hampering her ability to enjoy others, then that trait is a virtue.

For any candidate trait, the process of classification will be a two-part effort. First, we must settle on a characterization of the trait. In some cases this will be the challenging part, owing to confusions of usage and concept. By "modesty" do we intend the original meaning—sexual modesty—basically, a disposition to keep your clothes on and remain stern in the face of dirty jokes? Or do we intend the more recent usage—something closer to humility? If we intend the more recent usage, do we mean something superficial, like tact? Or do we mean something deeper?

After we are clear on the characterization of a trait, the second question is: what is this trait's typical effect on its possessor's relationships and projects? Traits that reliably benefit their possessor's relationships and projects are egoistic virtues; traits that reliably harm the same are vices.

5.2 A fragment of a catalog of virtues

In this chapter and the next, I gesture in the direction of a catalog of egoistic virtues. I give four examples that I hope, taken together, show the plausibility and the appeal of character egoism.

If ever there were a virtue that cried out for justification in egoistic terms, it is perseverance. It appears in Hume's catalog of virtues as a "selfish" virtue—that is, a virtue whose main benefit is to its possessor (EPM 6.21; SBN 243). So I begin with perseverance (and its vicious cousin ruthlessness) to show how to apply the theoretical account to the classification of a trait in an easy case.

Second, I apply the same method to self-respect. I dwell on self-respect for two reasons. First, self-respect is a centrally important concept in much moral philosophy, and character egoism can account for this centrality. Second, self-respect will play an important role in my discussion of justice and oppression in chapter seven.

Third, I discuss modesty. Modesty's status as a virtue has long been controversial, and I argue that evaluating modesty from the perspective of character egoism helps show the weaknesses of two different accounts of modesty as a virtue.

Finally, in chapter six I argue that character egoism is well-equipped to embrace honesty as an egoistic virtue.

5.2.1 Perseverance and ruthlessness

As with every candidate trait, before we can classify perseverance, we must first characterize it, then weigh its tendencies to promote or undermine relationships and projects.

Characterization. Perseverance, roughly, is the trait of character that urges us to continue pursuing a goal we have set for ourselves even when we meet obstacles.

Balancing benefits and harms. Most projects worth pursuing are intrinsically challenging. Learning an instrument, for example, is often frustrating and every beginner must endure a period of squeaks and screeches. Aphids sometimes invade our gardens. Our colleagues sometimes overlook our good work. Our legs sometimes give out on the trail. If we lack perseverance—if we tend to abandon projects when we meet frustrations like these—we are unlikely to succeed in any of the larger projects we care about. Whatever particular projects might be the ones for us, perseverance promotes success in them.

Does perseverance cause collateral damage to our ability to enjoy relationships, or to pursue some projects? Obsessive or monomaniacal focus on a project could harm our relationships and other projects. (Captain Ahab was a lonely man.) But obsession is not what we mean by “perseverance.” Perseverance is not a refusal to give up no matter what the cost. It is, rather, a disposition not to be easily deterred by obstacles. It isn’t clear how such a trait could harm any class of relationships, or any class of projects.

Classification. Perseverance is a virtue.

And what of ruthlessness?

Characterization. A ruthless person is one who is willing to do anything in pursuit of her goals, unbounded by feelings of sympathy or compassion for others. The ruthless person need not take an active interest in harming others, as does

the cruel person, but she is unperturbed at the prospect of harming others as a side-effect of pursuing her own goals, or as an effective means of achieving them.

Take as an example of ruthlessness the character of Eve Harrington from Mankiewicz' "All About Eve." At one point, Addison DeWitt, a theater critic, says the following. "You're an improbable person, Eve, but so am I. We have that in common. Also a contempt for humanity, an inability to love or be loved, insatiable ambition—and talent." Eve Harrington is willing to do anything to climb to the top of her profession. She feels no bonds of friendship or even mutual affection, she doesn't hesitate to lie to or betray those who have helped her. And her ruthlessness is effective. At the end of the movie, she gets what she covets: the Sarah Siddons Best Actress of the Year award.

Balancing benefits and harms. Ruthlessness can take us a long way toward achieving the goals we set for ourselves in our projects. In fact, we might see ruthlessness as something like perseverance on a steroid bender—not only is Eve committed to overcoming any obstacles thrown in the way of her acting career, she is committed to throwing up as many obstacles as she can in the paths of her competitors.

Ruthlessness seems likely to promote success most especially in those projects that require no cooperation and are competitive. However, ruthlessness is little advantage in projects that aren't competitive, and it is a terrible liability in projects that require cooperation. Worse, ruthlessness unfits us for most relationships. The ruthless person will frequently find herself friendless (and this is illustrated in the example of Eve). In fact, ruthlessness, because it entails both untrustworthiness and unkindness, seems to preclude almost every *close* relationship we can imagine. Indeed, to cultivate ruthlessness is often tantamount to *abandoning* relationships in favor of projects. This is not an appealing bargain.

Classification. As with smoking in the case of health, ruthlessness counts as a vice despite its beneficial side-effects. The downsides of ruthlessness are more certain and more devastating than its upsides.

5.2.2 Self-Respect

Self-respect is an important concept in much moral and political philosophy, especially in the Kantian tradition. Though it is developed at length and in many branches by different philosophers, there is fairly broad agreement on a minimal characterization of self-respect. At the least, I believe the various theorists of self-respect would accept the following claim: to have self-respect, a person must stand behind her own values, she must take herself to be the sort of person whose values matter, she must take herself to be the sort of being who is able to (and ought to) plot her own course in life.

Thus, Rawls: self-respect is “a person’s sense of his own value, his secure conviction that his conception of the good, his plan of life, is worth carrying out” (Rawls 1971, 440).

Thus, Dillon, writing about “basal self-respect,” the foundational form of self-respect, on which other layers of self-respect are built:

[Basal self-respect is] a prereflective, unarticulated, emotionally laden presuppositional interpretive framework, an implicit “seeing oneself as” or “taking oneself to be” that structures our explicit experiences of self and worth... When secure and positive, basal self-respect involves an implicit confidence, an abiding faith in the rightness of my being, the unexpressed and unquestioned (indeed, unquestionable) assumption that it is good that I am (Dillon 1997, 241-242).

Thus, Williams: a desire for self-respect is the “desire to be identified with what one is doing, to be able to realize purposes of one’s own, and not to be the instrument of another’s will unless one has willingly accepted such a role” (Williams 1973, 234).

More recently, we can see the same notion of self-respect behind Ronald Dworkin’s normatively loaded principles of self-respect and authenticity: “Each person must take his own life seriously: he must accept that it is a matter of importance that his life be a successful performance rather than a wasted opportunity.” And, further, every person has “a special, personal responsibility for identifying what counts as success in his own life; he has a personal responsibility to create that life through a coherent narrative or style that he himself endorses” (Dworkin 2011, 203-204).

Again, self-respect, as I use it here, is the disposition to stand behind one’s own values, to see them as worth having, or worth carrying out.

There are other concepts that are sometimes mistaken for (or confused with) self-respect.

Recognition, evaluation, and self-esteem

Some writers draw a distinction between *recognition* self-respect, and *evaluative* self-respect.¹ Recognition self-respect is what I’ve been discussing thus far. It is an attitude toward one’s status as a being. Evaluative self-respect is a judgment about how one is doing by one’s own values. For the reasons that follow, evaluative self-respect is not the relevant concept when discussing the *virtue* of self-respect.

First, note that recognition self-respect is a character trait, while evaluative self-respect is a judgment. Consider someone who says, “I always said I was the

¹ See Darwall 1977 for the foundation of this distinction.

sort of person who would never cheat on my wife. But now I've done it and I can't look at myself in the mirror anymore." This person's evaluative self-respect has taken a hit—he can no longer bear his own survey. But what has changed is his attitude toward (or beliefs about) his own character, not his character itself. (Though such a change might follow from his evaluation of himself as morally deficient.)

Second, note that low evaluative self-respect need not imply anything about a person's level of recognition self-respect. A robust and secure recognition self-respect is perfectly compatible with a low evaluation of oneself. In many cases, the cause of low evaluative self-respect can be found in changing beliefs or values.

Consider, for example, a white supremacist whose life is organized around earning glory in that community. If he comes, over time, to learn that the beliefs that ground his racism are false, his evaluative self-respect will surely suffer—that is, he will be ashamed of his former racist self—but this loss of evaluative self-respect doesn't entail a loss of recognition self-respect. That is, despite feeling ashamed of his former self, he may still consider himself to be the sort of being whose values matter. In fact, in this sort of case, it might be that recognition self-respect is a *precondition* of losing evaluative self-respect. In order to feel badly about using his valuing capacities to choose terrible things, the former white supremacist must believe he's someone whose values matter.²

Similarly, consider a man who, when young, valued material success and organized his life around succeeding in business. Later in life, a successful businessman, he comes to value other things. Quiet walks in the woods, maybe. Or Scrabble. These are not things he valued as a young man, and when he pursued those values, he was not operating under any special false beliefs. He changed. But from the

² The 1998 film *American History X* is the sort of example I have in mind.

perspective of age, the values of his former self might appear to him to be childish, or embarrassing, or shameful. He might have difficulty bearing his own survey, but this has nothing to do with an unwillingness to stand behind his own values.

Of course it's possible that low evaluative self-respect could be rooted in low or absent recognition self-respect. If I feel my own values weakly and am open to manipulation, I might find myself living someone else's values, and they may not match up well with my own. Imagine, for example, a student who majors pre-med and goes to med school because it's what her parents expect, even though she doesn't enjoy it and doesn't believe that's where her talents lie. She might reflect on her life, one day as a doctor, and conclude that she's done a lousy job of living her own values, because she was insufficiently willing to stand behind them.

Self-esteem is one's evaluation of oneself according to some standard. That standard might be one's own values (in which case self-esteem and evaluative self-respect are the same thing) but it need not be.

Consider Jones, who is sycophantically devoted to David Bowie. He adapts himself to Bowie's whim, becoming whatever Bowie wants at any moment. Jones sees Bowie as the appropriate setter of his values. He is servile; he lacks self-respect. But Jones esteems others and himself according to closeness to Bowie. Despite his servility, it is entirely possible that he can esteem himself highly, so long as Bowie showers more attention on him than other groupies. Jones could lack (recognition) self-respect and never bother to engage in the kind of reflection or introspection necessary to generate a judgment of evaluative self-respect, and nevertheless he could have high self-esteem.

As evaluations of self, neither self-esteem nor evaluative self-respect are the relevant concept for a discussion of the virtue of self-respect. When I speak of the *virtue* of self-respect, I am speaking of the character trait. That is, in the

parlance, *recognition* self-respect.

Why, then, is self-respect an egoistic virtue?

Self-respect as an egoistic virtue

Dillon, among others, makes welfare-based claims about self-respect's goodness:

Individuals who are blessed with a confident respect for themselves have something that is vital to living a satisfying, meaningful, flourishing life, while those condemned to live without it or with damaged or fragile self-respect are thereby condemned to live constricted, deformed, frustrating lives, cut off from possibilities for self-realization, self-fulfillment, and happiness (Dillon 1997, 226).

It is an advantage of character egoism that the justification for this sort of talk falls directly out of the definition of virtue. After all, from an egoistic perspective, self-respect is as close as is possible to a necessary condition of well-being: we can't hope to succeed in projects and relationships *of special importance to us* if we don't stand behind our own values. If we lack confidence in ourselves as valuing beings, if we allow other people to set our values, or organize our lives around what we believe the values of others to be, there's little chance that our life will be one that is good for us. Our lives may be good for those we serve, but not for us.

In short, if well-being is enjoying success in projects and relationships that are important to us, we can only hope to get off the ground if we are willing to stand behind our values—to be confident that we are the sort of people who ought to, who *deserve* to, pursue the projects and relationships that matter to us.

There are two potential worries about the egoistic approach that I think are misguided. The first is that character egoism will encourage the abandonment of self-respect if that serves some aim that is especially important to us. The second is that egoism about character will encourage the abandonment of self-respect in circumstances of oppression or injustice, in which maintaining self-respect is dangerous.

Strategic servility

When servility is cultivated, it is typically done in pursuit of a particular relationship. The servile person hopes to build or cement a relationship with his (for lack of a better word) master. This suggests a worry about the egoistic approach. If someone wants a relationship with a particular person more than anything else, and can secure that relationship via servility, wouldn't character egoism recommend servility instead of self-respect?

Orgon, the religious husband in Moliere's *Tartuffe*, is an example of strategic servility. In an effort to win hypocritical holy man Tartuffe's favor, Orgon signs all of his possessions over to him, offers him his daughter, Mariane, for a wife, and generally makes it clear that he conceives of himself as at Tartuffe's service. Orgon sees his own servility as a trait useful in promoting success in an especially important relationship. But his own case reveals the reasons servility is a lousy strategy.

First, Orgon's slavish devotion gets him nowhere with Tartuffe. This should come as no surprise: it's notoriously difficult to have a successful relationship with a doormat. Second, Orgon's servility to Tartuffe nearly destroys his relationship with his own family. This also should come as no surprise: while it's difficult to have a relationship with a doormat, it's nearly impossible to have a relationship

with someone else's doormat. Since servility entails the abandonment of personal projects, risks serious injury to other relationships, and is usually ineffective in securing the relationship it is intended to promote, it's a terrible strategy. It will not secure well-being.

But what if strategic servility *did* work? What if Tartuffe took on Orgon as a pupil? What if David Bowie settled down with sycophantic Jones? This still would not suggest that strategic servility can pay, because it is still the case that Orgon and Jones have restricted themselves to a single value—a single relationship. This slavish devotion to a single value is the mark of someone confused about his own well-being. It is rarely, if ever, the case that we can be happy this way. Orgon would probably be better off with hobbies and friends than with Tartuffe alone, for the simple reason that monomania is usually bad.

Given that strategic servility rarely works, and even if it did, such single-minded focus on a single value would only promote the well-being of an unusual person, it becomes difficult to press the challenge that egoistic virtue ethics risks endorsing strategic servility. To hit home, the challenge would have to be more like this: "if human beings could be happy pursuing a single value and were inclined to bond with sycophants—that is, if human beings were nothing like they are—then the egoist about character would endorse strategic servility as a virtue." This is true. But this form of the objection is very close to a simple restatement of a wide range of virtue ethics. For an Aristotelean virtue theorist, or a utilitarian virtue ethicist, or a character egoist, it is true that if human beings were not as they are—if human beings were dogs, for example—then we would endorse a different list of virtues, maybe even servility.

But human beings don't work that way, and the addition of strategy can't rescue servility from the catalog of vices.

Defensive servility

Another possible variety of objection: imagine an unjust, oppressive society. Suppose that for members of the oppressed group, living one's life according to one's own values—that is, living a life that manifests self-respect—risks pain, or imprisonment, or even death. Think of Kunta Kinte's plight in *Roots*. Wouldn't character egoism recommend that members of this group cultivate servility, to avoid these terrible outcomes? Surely such a recommendation is wrong on its face.

This objection might have real force if well-being were a matter of shallow hedonism. No doubt Kunta Kinte would have suffered less physical pain had he been servile and accommodating. But if we take seriously the *practical* account of well-being I've sketched, we should keep projects and relationships in mind. Is it the case that in some oppressive circumstances, people will more reliably succeed in projects and relationships that are important to them if they cultivate defensive servility than they will if they cultivate self-respect?

Given that self-respect is very nearly a *precondition* of having rewarding projects and relationships, circumstances that render self-respect a liability would have to be extreme, and the examples fantastical. Consider a situation of such terrible oppression, such *total* oppression, that resisting it—caring about any projects other than those dictated by the oppressors—reliably results in torture or death. People under this degree of oppression are more likely to succeed in projects of special importance to them without self-respect than with it. Thus, in this situation of hyperoppression, character egoism would recommend servility. At least, servile, you might be able to invest in and enjoy some aspects of the projects you've been forced to undertake. Take pride in the platinum effigy you're forced to build.

Maybe so. It isn't clear to me, though, that our intuitions rankle at this

suggestion. In circumstances of such awful oppression (and it is probably the case that we're talking about circumstances that have rarely existed in our world) in which maintaining self-respect is tantamount to suicide, it isn't clear to me that the agent who commits suicide-by-self-respect is praiseworthy. To justify such a claim would require, I think, that death is preferable to servility. Such a view is, at least, controversial.³

But let us remind ourselves that oppression this extreme *is* fantastical, which is why our intuitions are muddy.

In the real world, even terrible circumstances of oppression—American slavery, South African Apartheid, etc.—allow space for self-respect, even though they are (designed to be) hostile to it. Maintaining self-respect under these corrupt regimes is itself a victory, an accomplishment, which is part of why we admire those who manage it. The cost is high, but we pretty clearly believe the pain of maintaining self-respect is preferable to the cost of servility.⁴ And this should come as no surprise, because the cost of servility, on a view that holds projects and relationships as proxies for well-being, is tantamount to a total surrender of well-being.

All of which is to say that egoistic virtue ethics naturally accounts for the central importance of self-respect: it's very nearly a necessary condition of our own well-being, and it retains this importance through a broad range of circumstances.⁵

³ Would we recommend the suicidal course to our loved ones? Or would we council them to keep their heads down?

⁴ The high cost of maintaining self-respect in these regimes is part of what reveals them as corrupt. A legitimate political association ought to be structured around the effort to promote the well-being of the people who comprise it. Thus, a legitimate political authority should *foster* self-respect, not undermine it.

⁵ I discuss oppression and self-respect further in Chapter 7.1, where these issues come up again in slightly different context.

5.2.3 Modesty

Any plausible theory of the virtues must save the data. It must give an appealing account of why traits we consider bedrock virtues—traits like self-respect and honesty—belong in the catalog of virtues. Beyond this, we might hope a theory of the virtues would bring a new perspective to, or shed new light on, debates about controversial traits. Modesty is one such trait. Character egoism, for reasons that capture intuitions that have sometimes gone overlooked in the debate, does not classify modesty a virtue.

Character traits whose status is controversial might be controversial for two general reasons. It might be difficult to characterize the trait. That is, there might be confusion or disagreement about what disposition (or set of dispositions) the trait’s name indicates. Second, once the characterization of the trait is clear, we might disagree about whether or not that trait is virtuous or vicious or neither. Modesty is controversial along both these dimensions.

The set of behaviors associated with modest people is relatively uncontroversial. Modest people are graceful and generous winners. They don’t trumpet their own successes, and they’re comfortable acknowledging, frankly, the role other people, external circumstances, and luck played in those successes. They may be self-deprecating. Modest people are often uncomfortable receiving praise, especially effusive praise, but are happy to praise others.

Locating the trait of character (if there is one) that underlies and motivates these behaviors has proven a challenge. One complicating factor is that most of these behaviors could have underlying causes other than modesty; they could be caused by kindness or sensitivity, for example. Perhaps “modesty” is nothing but a shorthand way to refer to a collection of traits—kindness, sensitivity, tact, etc.—when those traits are manifested in the context of differential success. I’m

sensitive if I consider and accommodate your feelings when you're frustrated with your partner. I'm *modest* if I consider and accommodate your feelings when you're frustrated because I got the promotion you wanted. I'm *tactful* if I feel the urge to deflect attention from your loudly rumbling stomach. I'm *modest* if I feel the urge to deflect attention from the fact that I'm shellacking you at chess. And so on.

A rough analogy. There is a set of behaviors associated with show-boating. Show-boaters usually delay games, appealing for attention from the crowd. They're likely to taunt their opponents. They're prone to dancing. But these behaviors aren't usually thought to spring from an underlying character trait of show-boatitude. There are other traits that underlie the behavior: arrogance, over-competitiveness, insecurity, etc. When over-competitiveness manifests itself as ostentatious displays on the sports field, we call it show-boating. Similarly, it might be that when kindness manifests itself as self-deprecating behavior at an awards ceremony, we call it modesty.

With the promise that this issue will resurface shortly, I'll bracket it for now, and provisionally assume that there *is* a character trait of modesty that motivates the behaviors associated with the term. How, then, should we characterize modesty?

The understatement account rejected.

Perhaps modesty is the disposition to *understate* our view of ourselves. When a modest person thinks she's super great at something, she says things that suggest to the people around her that she thinks she's only OK. She understates her greatness.

Something like this seems to be Hume's view:

He must be a very superficial thinker, who imagines, that all instances of mutual deference are to be understood in earnest, and that a man would be more esteemable for being ignorant of his own merits and accomplishments. A small bias towards modesty, even in the internal sentiment, is favourably regarded, especially in young people; and a strong bias is required, in the outward behaviour (EPM 8.10; SBN 264-265).

Hume's account, on which modesty is the behavior of soft-pedaling one's greatness in order to spare the feelings of others, saves the appearances. Someone disposed to understate their view of themselves will be inclined to display the behaviors we associate with modesty: self-deprecation, attention-deflecting, and so on. It also gives an account of why the trait is a virtue: it helps us avoid making ourselves obnoxious to others. If we are frank about our view of self, we risk behaving vainly: displaying our awards, accepting (or worse, seeking) praise. These vain behaviors, Hume says, consist "in such an importunate and open demand of praise and admiration, as is offensive to others, and encroaches too far on *their* secret vanity and ambition" (EPM 8.11; SBN 266). Modesty helps us avoid these pitfalls.

Though Hume's version of modesty characterizes the trait in a way that saves the appearances and explains why the trait is good, there's a problem: there's another word for the behavior of thinking oneself great but saying otherwise. As Julia Driver says, people "who merely understate their self-worth, even though they are fully aware of it... exhibit false modesty" (Driver 2001, 17). False modesty might still be a virtue—Hume makes a pretty good case!—but that question is distinct from the question of modesty. The understatement account of modesty, then, characterizes the wrong trait.

Two competing views: Underestimation (Driver)

Julia Driver believes that some virtues analytically include ignorance. Modesty is the flagship in her fleet of virtues of ignorance. According to Driver, a modest person will not only *say* self-deprecating things about herself, but will *believe* them. “An agent is modest if he is disposed to underestimate self-worth to some limited extent, even in spite of the available evidence” (Driver 1999, 830).

In typical cases, someone who underestimates her self-worth is likely to display the behaviors we take to mark modesty. She’ll probably be self-deprecating (because she believes her accomplishments aren’t worth as much as other people think they are) and she’ll be likely to deflect attention away from herself (because she doesn’t believe she deserves it) and so on.

Driver’s account of modesty’s goodness is the same as Hume’s account of (false) modesty’s goodness. She believes modesty is a virtue because it lubricates social interactions. Someone who undervalues herself, and thus behaves modestly, “seems less likely to provoke an envy response in others” (Driver 1999, 828). Happier social interactions benefit everyone involved, and in that benefit lies modesty’s goodness.

Two competing views: Not caring (Schueler)

G.F. Schueler has a competing account of modesty.⁶ He suggests that “someone who is genuinely modest is someone who does not care whether people are impressed with her for her accomplishments (where ‘accomplishments’ should be

⁶ Schueler rejects Driver’s account on several grounds. Chief among them is that her definition *doesn’t* save the appearances: “Someone could be disposed to underestimate her own worth, due to some systematic, minor miscalculation, say, and yet still be a full-blown, loud-mouthed, braggart, that is, a paradigm of someone who is definitely not modest” (Schueler 1999, 835). This is potentially a knock-down objection. In order to illustrate character egoism’s contribution, though, we need only Driver’s and Schueler’s positive views.

read very broadly, so as to include possessions, abilities, etc., anything about which one might be or fail to be modest)” (Schueler 1999, 838). The claim is not that modest people are disdainful of the opinions of others; modest people are *indifferent* to others’ evaluations of their accomplishments.

Schueler’s definition saves the appearances because the behaviors we associate with immodesty—certainly the most egregious ones—are clearly intended to secure attention and praise. Gloating, bragging, lording-over—these behaviors only make sense if the person doing them cares a great deal about impressing others.

Schueler’s account of the goodness of modesty diverges widely from the tack Driver and Hume take. The person who cares about impressing others, Schueler says, “is the sort of person whose direction in life, whose goals and purposes and so on, are generated not from herself but from those around her” (Schueler 1999, 838). Conversely, modesty is a virtue “because of what it reveals about the person who has it, namely, that her goals and purposes come from herself, not from others. Someone who is genuinely modest is thus seen to have a kind of substance to her character, just the sort of substance that the immodest person lacks” (Schueler 1999, 839).

Competing views evaluated

The perspective of character egoism sets in relief problems with both these views, and gives some structure to the intuitive discomfort we might feel when considering Schueler’s and (especially) Driver’s accounts of modesty.

When we take the egoistic perspective, it quickly becomes clear that Driver’s definition of modesty is open to a sensible-knave-type objection.⁷ So much the

⁷ The sensible knave, in Hume’s *Enquiry Concerning the Principles of Morals*, objects that all of the benefits Hume attributes to honesty are achievable via undetected dishonesty. At base, it’s an objection that Hume hasn’t located any benefit in the virtue, but rather has located all

worse for Driver, it's a knave-style objection that doesn't require knavery.⁸

Note that the ignorance Driver thinks constitutes modesty comes with real costs. If I underestimate my talents and abilities in those areas in which I am most talented and able, I pay a price in my ability to pursue projects in those realms. If I'm ignorant of my abilities I will confine myself within imaginary boundaries. Thus Driver's version of modesty reduces the chances that we'll succeed in the projects that are of special importance to us. What is best for us, if we want to succeed in the things we care about, is to have an understanding of our *real* potential and our *real* limits.

For Driver, the goodness of modesty lies in its promotion of healthy social relationships. This is a benefit that deserves consideration, and it is a benefit to relationships that could, conceivably, outweigh its cost to our projects. What seems glaringly clear, though, is that the social benefits that follow on underestimating ourselves can easily be had in other ways.

In fact, one might suspect that an accurate evaluation of our self-worth is more than enough, in most cases, to avoid the interpersonal pitfalls that arrogance presents. When we complain about braggarts, it is usually in cases of people boasting *without warrant*. If a good swimmer I meet at the YMCA claims he's the best swimmer in town, I might be put off by his boastfulness. If Michael Phelps tells me he's the best swimmer in town, he's probably right. In fact, in Phelps's case, if he tried to claim (even in honest ignorance) that his consistent success is undeserved because there are better swimmers out there, his underestimation of

the benefit in the *appearance* of having the virtue.

⁸ Schueler notes that undetected false modesty brings all the social benefits that Driver's modesty brings (Schueler 1999, 837). This is a direct analog of Hume's sensible knave, applied to modesty instead of honesty. I will go on to develop an analogous objection that doesn't require deception. That is, a way to appear modest that doesn't require either ignorance or deception.

himself might well be off-putting.⁹

Even in situations in which an honest evaluation of our abilities would alienate our peers—and even if these situations are common—we need not go nuclear and adopt Driver’s underestimation account to avoid bad consequences. A little bit of grace, or sensitivity, or empathy, or tact will do the trick. In short, Driver’s modesty is a poor candidate for an egoistic virtue because the trait’s benefits can be had without incurring its costs.¹⁰

Our pre-theoretic intuitions match character egoism much better than Driver-underestimation. We do not, for example, raise our children to be Driver-modest. No doubt we don’t want our children to be *vain*. But would minimally decent parents choose to raise their children in such a way that the children had artificially and mistakenly low views of themselves? Parents could insure this outcome by denigrating their children, especially in their areas of special talent. But we’d condemn this approach to parenting even if (especially if) it achieved the goal of undermining the child’s self-image. Surely a child who is exceptionally talented in some area should understand this about herself. It’s a failure of parenting if our children, particularly talented in some area, view themselves as merely typical in that area. We don’t want people we care about to underestimate their own self-worth, because it’s bad *for them* if they do.

Schueler’s case is different. His account of the goodness of modesty is grounded

⁹ Of course there are circumstances in which Michael Phelps could make himself obnoxious to others via a frank assessment of his own ability. Poolside before a race, or at a restaurant with his competitors after he’s just won, Phelps’s honest assessment of himself as the best in the world might be inappropriate.

¹⁰ As always, with character egoism, it’s tough to make categorical claims. There might be unusual people who are incapable, for whatever reason, of acting gracefully, or tactfully, or with sensitivity, whose accurate estimation of their own abilities would make them obnoxious to others. For these people, ignorant modesty might allow them to form relationships, and this benefit might outweigh the costs to their projects. But that situation seems rare—if not fantastic—and for most of us, who have basic social skills and moderate abilities, ignorant modesty is a pointless sacrifice of welfare.

in the agent's attitude toward her own values. "The immodest person, or at least the paradigmatically immodest person at the very far end of the continua... will be someone who has no goals or purposes of her own at all, someone for whom all of her direction in life comes from others" (Schueler 1999, 839). This is a terrible outcome, in any case, and it's certainly terrible from the perspective of character egoism. If our goals and purposes are set for us by others, it is unlikely that our relationships and projects will be of special value to us. Character egoism will judge Schueler-modesty not just a virtue, but a centrally important virtue.

This should serve to focus our attention on the strangeness of Schueler's characterization of modesty. The fully immodest person, on his view, "has no goals or purposes of her own." But this is—uncontroversially, I think—also a definition of *servility*. Conversely in the case of a modest person, "her goals and purposes come from herself, not from others" (Schueler 1999, 839). This is very close to what I suggested, above, as a minimal characterization of *self-respect*.

What Schueler has noticed is that in the real world, self-respect is challenged in other ways than direct assaults. In the swirl of our social influences, the appeal of praise and admiration can be a strong enough pull that it encourages us to undertake projects and relationships in order to receive praise and admiration, even when those projects and relationships aren't of special importance to us. We can (and should) cultivate the kind of self-respect that allows us to stand in this swirl of influences without losing sight of what's important to us. We need the ability not to care—or at least, not to care too much—whether other people are impressed with our relationships and projects. Schueler terms this modesty. It is hard to see, though, how the trait he characterizes is anything other than self-respect in a particular social setting. Much like show-boating is over-competitiveness manifested as ostentatious display on the sports-field, Schueler's modesty is self-respect

manifested as indifference toward social evaluations of personal impressiveness. Neither show-boating nor Schueler-modesty seem like good candidates for their own, separate virtue.

The upshot

Adopting the perspective of character egoism helps organize and explain some of the controversies surrounding modesty's status as a virtue. Driver believes modesty is a virtue because being ignorant of one's own merit supports healthy social bonds. That is, Driver believes there's an egoistic reason for underestimating oneself. But we can quickly see that this is not so—the social advantages of underestimation are available to us without having to bear the costs of ignorance. Driver thinks she has identified advantage in modesty, but she is wrong.

Schueler believes modesty is a virtue because not caring about the evaluations of others frees us to set our own agenda, instead of chasing after the agendas of others. That is, Schueler believes there's an egoistic reason for not caring about the evaluation of others. But, phrased this way, we can quickly see that the genuine virtue Schueler sketches is the same as the familiar characterization of self-respect. Schueler thinks he has identified advantage in modesty, but he has actually identified advantage in something else.

In short, character egoism helps advance the debate about modesty. Modesty is probably not a virtue; it is commonly thought to be one because people mistakenly see advantage in it.

5.3 Conclusion

The way to classify a character trait is to consider, as carefully as we can, the role that trait plays in those relationships and projects that are (or ought to be, for our own sakes) important to us. Traits that unfit us for relationships or undermine our ability to engage projects are vices. Traits that promote the same are virtues.

The samples I've given in this chapter barely scratch the surface of a complete catalog of virtues. Nevertheless, I hope it's clear how the method I've sketched here would classify many traditional virtues as virtues. I think it is not controversial to claim that familiar virtues like courage, temperance, generosity, and kindness are traits of character that promote healthy relationships and rewarding projects.

There are, however, two familiar virtues that many people believe are irreducibly other-directed: honesty and justice. These two virtues are the subject of the remainder of the dissertation.

Chapter 6

Egoistic Honesty: Responding to the Sensible Knave

One must be fond of people and trust them if one is not to make a mess of life.

“What I Believe”

EM FORSTER

Historically, egoists and their sympathizers have struggled to account for the virtue of honesty. The *prima facie* case that honesty benefits others—but not its possessor—is strong: in situations in which we are tempted to lie, it’s almost always because we believe things will go better for us if we do. We *sacrifice* something when we are honest—when we admit we’ve been given too much change, when we own up to a mistake we could have passed off on someone else, when we confess infidelities, and so on. Usually, when we choose honesty—when we choose the apparently self-sacrificing course—we think of honesty as something we owe the person we’re being honest with. I could walk away with the extra change in my pocket, but I feel I owe it to the careless clerk to mention the mistake. I feel

I'd be wronging him if I feign ignorance and keep the money.

There are certainly social benefits that follow on a *reputation* for honesty. My business endeavors will go better if my colleagues believe I'm honest. My romantic relationships will go better if my partner believes I'm honest. If genuine honesty were the only way to secure a reputation for honesty, there would be a straightforwardly egoistic case for honesty as a virtue. Sure, the liar may benefit from the extra change he pockets, but his reputation for dishonesty hampers all his relationships, and his socially engaged projects. Better for us all to forego the petty profit, and instead enjoy the much more important benefits of honesty.

Why should we believe, though, that genuine honesty is the only way to achieve a reputation for honesty? We've all known people who could lie, undetected, for a long time. How can we be sure that, of the people we believe to be honest, we will not discover, some day, that they aren't? Maybe there are even people we believe to be honest who will keep us fooled forever.

Because the egoistic advantages of honesty are usually assumed to flow from a *reputation* for honesty, and a reputation for honesty is not reliably linked to the character trait of honesty, honesty remains a problem virtue for egoists. It seems the most beneficial trait is not genuine honesty, but rather *apparent* honesty, conjoined with secret, opportunistic, dishonesty. But this is an unacceptable trait for a catalog of virtues—it flies in the face of settled intuitions about honesty's status as a virtue.

This problem is outlined by Hume at the end of the *Enquiry Concerning the Principles of Morals*. The Sensible Knave claims that the egoistic view of character can't capture honesty's goodness. Instead the real egoist—the Sensible Knave—should cultivate a disposition to tell the truth only when that benefits him, but to lie undetected when that benefits him.

There is fairly broad agreement that Hume can't answer the Sensible Knave, and that this reveals a deep problem in his moral theory. It is true that Hume *doesn't* answer the Sensible Knave. But I believe his moral theory includes the tools needed to answer him. In this chapter I discuss how Hume can answer the Knave, and I embrace this answer for my own theory as well. Hume's best answer to the Sensible Knave shows how honesty—genuine honesty—benefits its possessor more than secret, successful, and profitable dishonesty. And this, in turn, shows why honesty deserves a place in an egoistic catalog of virtues.

6.1 Hume's problem

Hume maintains that it is a central strength of his moral system, over all those that came before it, that it can show in “particular detail, that all the duties, which it recommends, are also the true interest of each individual” (EPM 9.16; SBN 280). Noting this, the sensible knave stalks Hume's theory of morality with the following question: ought I not, then, act morally in those cases in which I might be found out, but do whatever adds to my fortune, whether honestly or dishonestly, in those cases in which I cannot be caught?

There is no dearth of debate in the secondary literature about whether or not Hume can answer the knave's challenge. Out of this debate seems to have emerged a rough consensus that Hume cannot meet, head on, the challenge of the sensible knave.¹

¹ A few of the important discussions on which I will not comment, that make up this consensus, are:

- Gauthier 1992. Gauthier maintains that any answer Hume tries to offer the sensible knave will lead him to an error theory. That is, Hume will have to confess that all our judgments about justice are predicated on the mistaken claim that we have duties to justice in the first place. He attributes a similar view to:
- Mackie 1980.

Of the papers that constitute this rough consensus the closest, in substance, to the answer I propose is Gerald Postema's "Hume's Reply to the Sensible Knave." Postema, after a careful reading of the relevant passages in the *Treatise* and *Enquiry*, uses the tools of Humean psychology to construct what he takes to be Hume's strongest answer. He then finds this answer lacking. However, there are exegetical reasons to believe that my reading more closely approaches the kind of response that Hume had in mind, and there are philosophical reasons to prefer my reconstruction of Hume's response to Postema's.

6.1.1 The knave's challenge

Hume insists that justice is a hard and fast rule. Clearly, acts of justice will often be contrary to private interests, and Hume further grants that "a single act of justice is frequently contrary to *public interest*." But even in these situations, the utility derived from observing the rule as exceptionless outweighs the utility derived from acting for public or private utility on a case-by-case basis. "'Tis impossible to separate the good from the ill," says Hume. Thus, and because "without justice, society must immediately dissolve," we must pursue "the steady prosecution of the rule" of justice in all cases and without exception (T 3.2.2.22; SBN 497).

The knave's objection is straightforward.

A sensible knave, in particular incidents, may think that an act of

-
- Baron 1982. Baron claims that the truth about justice is pernicious, and so Hume constructs a noble lie intended to ensure the continued functioning of society.
 - Stroud 1977. Stroud concludes his discussion of the sensible knave by saying "If there are any occasions on which a man reasonably thinks he can get away with injustice and profit from it then there will be nothing to recommend justice to him at all" (210).

iniquity or infidelity will make a considerable addition to his fortune, without causing any considerable breach in the social union and confederacy. That *honesty is the best policy*, may be a good general rule, but is liable to many exceptions; and he, it may perhaps be thought, conducts himself with most wisdom, who observes the general rule, and takes advantage of all the exceptions. (EPM 9.22; SBN 282-283)

What the knave notices is that, for society to enjoy the fruits of justice, all that is required is the *appearance* of justice. In terms of public utility, a fiction of justice, universally believed, is better than a fact of justice, commonly doubted. This opens up an opportunity for the knave. If he can take advantage of those situations in which he can unjustly pursue his own interests without perturbing the public perception of justice, then he has privately gained while doing no public harm. In fact, were he to pass up these same opportunities, he would bestow no benefit on society; he would merely *fail* to benefit himself. If morality is coincident with the true interests of the individual, as Hume suggests, then how can we avoid the conclusion that the knave ought to pursue dishonestly his own interests in those situations in which he is not detectably harming anyone else's?

Hume is adamant that, as a practical matter, the knave's is bad advice. It will always be the case, human nature being what it is, that a knave will go for one heist too many. Knaves will succumb to greed and be found out, suffering "a total loss of reputation, and the forfeiture of all future trust and confidence with mankind" (EPM 9.24; SBN 283). It is important not to mistake this argument for the one that engages the philosophical thrust of the knave's challenge. The philosophical question is: *were* there an entirely "secret and successful" knave, ought she or ought she not to act justly?

A very, very good liar does not, in our esteem, compare favorably with an

honest person. We wouldn't approve of dishonesty, or counsel it for those about whom we care, even if we were confident in their ability to pull off a life-long, flawless performance. Why? What justifies the moral intuition that honesty, and not merely the appearance of honesty, is good?

To be clear, the most challenging knave, the most "secret and successful" knave, is not merely one whose misdeeds cannot be traced back to her. It is a knave whose misdeeds are never noticed at all. This is the limiting case of those unjust actions that can be accomplished "without causing any considerable breach in the social union and confederacy." This is the form of the knave's challenge that requires an answer.

Hume devotes the closing paragraph of the *Enquiry* to the knave's challenge. But to unpack that answer that lies in that paragraph is not an easy task.

6.1.2 Postema's reconstruction of Hume's response

Because Hume maintains that all our moral approbation has its source in utility or agreeableness to the agent or others, Hume must meet the challenge of the knave by showing that there is disutility (or disagreeableness) in knavery—either to the knave himself or to others. Postema recognizes that the challenge is set up so that there can be no public disutility—after all, the knave's injustices go unnoticed. If we are to find disutility in knavery, we must find it within the knave himself.

Postema discovers possible damage to the knave after analyzing Hume's account of the social basis of character. He notes that for Hume, in order to discover one's own character, one needs "reflection and confirmation of one's soul in the souls of others." Because the life of a successful knave requires that he constantly and effectively hide his true motivating principles, the knave will be unable to

enjoy this reflection, and thus will find himself without a character. “The ‘successful’ knave,” Postema says, “gains the whole world only to lose his own soul. And this, says Hume, is a bad bargain for any person” (Postema 1988, 35).

This reading does engage the philosophical thrust of the knave’s challenge, but it is not problem-free. Postema himself brings the first objection. It is clear that, to develop a character, a man need not be reflected in the souls of all humanity. In fact, it might be sufficient to be reflected only in the eyes of a family, or a close circle of friends. The knave, then, might be able to expose his knavery to some small group, and in exposing his true motives for reflection in their souls, develop a character. Thus can the knave both have and eat his dastardly cake, secretly exploiting most people, but gleaning the benefits of a character from the small group with whom he is honest. Postema thinks this is a knockdown objection, and concludes that Hume’s reply to the sensible knave is “ultimately unsuccessful” (Postema 1988, 23).

I think this objection, as presented by Postema, misses a deeper problem. Hume does have an account of character generation that is irreducibly *social*. But the concept of character for which Hume offers this irreducibly social ontogeny is a concept we hold in common independently of Hume’s account of its origin. A character is something like a unity or consistency of values and motives that provides the stable center of a person. Someone has *no* character if they lack this consistency of values and motives. But if we look for a unity of values and motives in the knave, we find it in spades. When we describe a knave as someone with an inordinate desire for property, who carefully maintains a facade of honesty and justice while pursuing every profitable opportunity for dishonesty and secret injustice, we are describing the knave’s *character*.

Postema’s reconstruction of Hume’s response runs into trouble, then, because

it concludes something—namely, that the knave is left without a character—that is contradicted by the framing of the challenge. While it is certainly fair to say that the knave has a *lousy* character, it seems flat wrong to say that he has no character at all.

Fortunately, for the defenders of Hume, there are some textual indications that Postema’s response is not the response that Hume intended.

Hume prefaces his response to the knave by saying that it is unlikely that anyone inclined to knavery will be convinced by his response. “If a man think that this [knavish] reasoning much requires an answer, it would be a little difficult to find any which will to him appear satisfactory and convincing” (EPM 9.23; SBN 283). With this in mind, Postema’s reply seems almost *too* good.

If the result of knavery is the annihilation (or preemption) of character, then the knave will lack consistency in his values and motives. If he lacks this stability, he will act in ways that subvert his future aims and contradict his past values. In short, he will continually frustrate his own will. To develop a frustrated will is contrary to self-interest, even understood in the narrowest, money-making sense. As someone who is overly moved toward material gain, this is a state of affairs the knave will want to avoid. Thus, if the annihilation of character can be shown to be the result of knavery, as Postema seems to have done, it is hard to believe that this line of thinking will *not* convince the knave. Since Hume says that the knave will remain unconvinced, this is a flag that Postema has not reconstructed the response as Hume would have it.

Much of the motivation for Postema’s character-oriented reading seems to come from the first sentence of the last paragraph of the *Enquiry*:

But were [knaves] ever so secret and successful, the honest man, if he has any tincture of philosophy, or even common observation and

reflection, will discover that they themselves are, in the end, the greatest dupes, and have sacrificed the invaluable enjoyment of a character, with themselves at least, for the acquisition of worthless toys and gewgaws (EPM 9.25; SBN 283).

Postema reads “have sacrificed the invaluable enjoyment of a character” to mean “have sacrificed their characters.” This is a puzzling gloss, since there are at least two ways to sacrifice the enjoyment of character. One could, of course, entirely lack a character. But one could also find oneself unable to let one’s character interface with the world in the ways required to *enjoy* that character.

And that brings me to my suggestion for how to interpret Hume’s response to the sensible knave.

6.2 Trust, friendship, and knavery

The problem with knavery lies not in its effects on something the knave is unable to collect from the community (character), but in its effects on something a normally functioning person extends to others: trust.

Regardless of how clever the knave may be, he can never escape the fact that he knows he is a knave. His entire *modus operandus* turns on his awareness of the space between the secret maxim that governs his actions and the maxim he publicly promulgates but chooses not to follow. He knows that he is hiding dishonesty behind a veil of honesty.

If the knave is aware of his own deceptions (and also aware of how effective this deception is), then might he not also suspect that others in the world are practicing these same deceptions? This sort of character, the man with the secret vice who is constantly suspicious of the same vice in others, makes intuitive sense.

It is certainly a character well represented in history and literature.

A good literary example is Iago, the villain of *Othello*. Iago is on the short-list of the most treacherous, dishonest characters Shakespeare wrote, and yet he's so effective in his knavery that the other characters in the play refer to him, un-ironically, as "Honest Iago."

Early in the play, Iago addresses the audience to explain why he hates Othello. His main reason is that he suspects Othello is sleeping with his wife, Emilia (Shakespeare 2001, 1.3.379). It isn't true, of course. But what makes this moment interesting is that Iago has *absolutely no evidence whatsoever* of infidelity on the part of his wife. Othello at least has circumstantial evidence that Desdemona is cheating on him. Iago has not the slightest hint of any kind that Emilia is cheating on him.

It's hard to know how to read Iago's mistrust of his wife without appealing to his own constant treacheries. That is, it is *because* he's constantly cheating on her, while *appearing* to be faithful to her, that Iago begins to count Emilia's *apparent* faithfulness to him as evidence of her infidelity. Because of his dishonesty, Iago has lost the ability to trust Emilia.

Iago's deceit-bred paranoia seems plausible enough. But does *Hume* have an account of how it might come to pass that someone will, through secret treachery, become suspicious of betrayal in others?

In fact, Humean psychology explains this phenomenon very naturally. The knave will have an impression of his own methods and motives, gleaned directly from his experience of them. He will also have an impression of his external appearance; that is, he will have an impression of how he appears to others. After all, "the minds of men are mirrors to one another" (T 2.2.6.21; SBN 365). These two impressions will be associated due to constant conjunction, as he is the original

source of both impressions. Finally, the knave will receive an impression of those persons who surround him, persons who are, apparently acting honestly. We would expect the knave's mind to associate the impression of apparently honest action from others with the impression of his own apparently honest action, and then associate the honest appearance of others with his own dishonest motives. To explain Iago in Humean terms requires only the tendency of the mind to associate resembling and constantly conjoined impressions.

Such knaves, in a state of general paranoia about the motives and methods of others, will find that "their selfishness and confin'd generosity, acting at their liberty, totally incapacitate them for society" (T 3.2.2.24; SBN 498). The knave will find himself gaining material rewards from his dishonest manipulation of society, but, thanks to his inability to trust others, he will be unable to enjoy its basic and more important pleasures: things like camaraderie and conversation.

Hume concludes the *Enquiry* with a meditation on the value of those human goods that have nothing to do with property.

And in a view to *pleasure*, what comparison between the unbought satisfaction of conversation, society, study, even health and the common beauties of nature, but above all the peaceful reflection on one's own conduct: What comparison, I say, between these, and the feverish, empty amusements of luxury and expence? These natural pleasures, indeed, are really without price; both because they are below all price in their attainment, and above it in their enjoyment (EPM 9.25; SBN 283-284).

Because the knave trades the first two of these priceless goods, conversation and society, for "empty amusements," he is engaging in a losing bargain. Thus it is that the true interest of the individual lies in eschewing knavery.

The argument goes like so. Effective knavery will leave the knave unable to trust his peers. An inability to trust is tantamount to an inability to have meaningful friendships. Friendship is a more important component of a happy life than are the material gains the knave enjoys. Therefore, knavery is a losing bargain. When Hume says that knaves “have sacrificed the invaluable enjoyment of a character,” this is just what he means. Knavery renders the sensible knave unable to interface with others in a way that is fundamental to a happy human life.

6.2.1 An exegetical advantage: the knave isn’t convinced

To reconstruct Hume’s response in this way has an interpretive advantage over Postema’s reconstruction. As noted earlier, Hume believes that, if someone is inclined to knavery, “it will be a little difficult to find any [answer], which will to him appear satisfactory and convincing” (EPM 155.23; SBN 283). But Postema’s claim that knavery *destroys* character would convince the knave, were it true. This is because an absence of character will result not only in an inability to enjoy the immaterial goods of society, but also in an inability to carry out the coherent plans that are the hallmark of the unified values and motives that constitute a character. The inability to carry out coherent plans will have *material* ramifications—the character-less knave will frustrate his own ability to get the things he wants out of life. Thus, Postema’s response seems to speak to the knave on his own terms. And if the knave can be shown he might get fewer goods by acting knavishly than he will by acting honestly, his motivation to act knavishly will be undermined.

If, rather than losing his character, the knave loses his ability to trust, his knavish motivation will remain untouched. Someone inclined to knavery places a disproportionately high value on material gain. An inability to enjoy friendships

or engage in meaningful social interaction will not affect the knave's ability to profit from his knavery—after all, his peers will still trust him. Hume's response, as I read it, is to say “yes, you'll benefit materially, but you will sacrifice goods that are more important than material gain.” It is not surprising if the knave, who already excessively desires material gain, will remain skeptical. After all, the knave has his toys and gew-gaws to show for his knavery. Hume can offer no such concrete evidence of the goods that follow upon conversation and society. The knave, predisposed to think that property is more important than it really is, will not be convinced by a response that says that property is not as important as he thinks.

6.2.2 An objection: But the knave isn't convinced?

One might be inclined to see the *exegetical* appeal of this reading—that it preserves Hume's claim that the sensible knave will not be convinced by his response—as a *philosophical* weakness. After all, how interesting a response can it be if it cannot be deployed against the very people who demand an answer? Indeed, commentators have taken note of Hume's comment that “it would be difficult to find any” answer to the knave and made hay.²

I think this objection is not the chugging juggernaut it might initially appear to be. In general, the commentators inclined to bring this objection are making wrong assumptions about what criteria will support a successful response to the knave. And once again, there are exegetical reasons to believe that this was Hume's own position, and philosophical reasons to think he is right.

² Marcia Baron in “Hume's Noble Lie,” e.g., quotes this passage, saying “Hume's reply strongly supports my interpretation” (Baron 1982, 553). Her interpretation is that the morality Hume presents is a noble lie, intended not to be true, but to produce the norms most conducive to a well-functioning society.

Hume's comment about his inability to satisfy the knave is not some revealing moment of candor, much less an implicit admission of his theory's failure. It is, instead, a comment on the character of the knave himself—a comment that Hume had already framed in the opening paragraphs of the Enquiry.

With the exception of disingenuous disputants, says Hume, “disputes with men, pertinaciously obstinate in their principles, are, of all others, the most irksome.” The origin of the frustration in dealing with these two types is that “as reasoning is not the source, whence either disputant derives his tenets; it is in vain to expect, that any logic, which speaks not to the affections, will ever engage him to embrace sounder principles” (EPM 1.1; SBN 169).

Any knave who vigorously defended knavery would fall into the category of the pertinaciously obstinate.³ The position she has staked out—that material gain is more important than meaningful social interaction—is simply not one that can be argued against.⁴ Because the commitment to knavery does not have its source in reason, but rather in a lack of experience, the knave cannot be met with philosophical arguments. The knave's mind will only change when she has an experience of genuine friendship, and then can weigh anew her belief that material gain is more important.

In recommending against knavery, Hume's intended audience is not knaves but rather those who are something less than pertinaciously committed to knavery. This includes those of us who are academically interested in knavery, those of us

³ To the extent that we believe it is impossible for someone to honestly maintain that material gains are a surer and richer source of happiness than friendship, to that extent we may suspect that a sensible knave is a disingenuous disputant. But here we will give the knave the benefit of the doubt and assume she actually believes what she says.

⁴ Imagine coming upon someone who insists that she is happier when she is unhealthy than when she is healthy. What might one say in response? We know from experience, not argumentation, that, other things being equal, good health is better than illness. And yet this is just what she denies. The case is similar with knavery.

tempted to knavery, those of us who are at a loss to respond to knavery when we come upon it. But an answer that is compelling to these audiences will fail to move the knave, because the response depends on the premise that healthy relationships are more important than toys and gew-gaws. So be it. This is hardly more a philosophical weakness of Hume's ethical theory than it is a weakness of symbolic logic that my pertinaciously obstinate friend refuses to believe that *modus ponens* is a valid rule of inference.

6.2.3 Philosophical advantages

While the trust-based response seems closer to Hume's intent than Postema's character-based response, it is also more difficult to assail on philosophical grounds. Remember that Postema's account depends on the knave sacrificing his character through the secrecy inherent in knavery. However, the knave can be open with a small group of people, thereby maintaining his character, and secretive with everyone else, thereby maintaining his knavery.

If the problem with knavery is the erosion of the knave's ability to trust, this sort of challenge cannot be brought. The problem for the knave is that his mind associates the appearance of honesty with dishonesty. So long as the knave consistently takes advantage of opportunities for injustice while maintaining the affect of justice, this connection will be strong. Even in the midst of a close circle of friends, the knave will be suspicious of the characters of others. On the "trust" reading, there is no obvious loophole through which the knave can enjoy the fruits of his knavery without suffering the consequences, because the consequences for knavery flow entirely from knavery itself.

Finally, this reading of Hume has the advantage of painting a plausible psychological portrait of the knave. As elegant as Postema's reading is in other respects,

it simply seems wrong to claim that the knave lacks a character. As someone who is consistently motivated by the desire to increase his property by any expedient means, the knave appears as someone with a *bad* character. And this is the understanding of knavery that the “trust” reading invites.

It seems, then, that Hume has a viable response even to the ultimate formulation of the knave’s challenge. No matter how effective the knave, he will be unable to trust the other members of his community. In such a situation, the knave is denied the most basic and important goods of society. In this way, Hume shows that knavery is not in our self-interest.

6.2.4 Empirical advantages

Humean psychology suggests that dishonesty corrupts an agent’s ability to trust. Before character egoism can embrace Hume’s answer to the knave, we must have some reason to believe Humean psychology is *right* on this point.

As I noted before, using the example of Iago, there are many characters in stories and plays who display both dishonesty and corrupted trust. Folk psychology also seems to embrace this mechanism. In fact, it’s enshrined in the standard advice that “if he irrationally accuses you of cheating, it’s because *he’s* cheating.” These kinds of evidence give us some reason to believe the relevant psychological mechanism is real.

Recently, positive psychologists have gathered evidence that appears to corroborate this tenet of folk-psychology. Most importantly, researchers have found that a crucial aspect of friendship-building is *self-disclosure*. Thus, Argyle:

Close friends are a particular source of happiness. Weiss found that to avoid loneliness people needed a single close relationship and also

a network of relationships. To form a close relationship involves an increasing level of self-disclosure, and without it people will still be lonely (Argyle 2001, 75).

But self-disclosure is unavailable to the knave, who deliberately seeks a gap between his reputation and himself. Thus do literary examples, folk psychology, and positive psychology all suggest that those, like the knave, who lack the ability to self-disclose will struggle to form close relationships.

6.3 Conclusion

A *reputation* for honesty is important for the success of our socially engaged projects, and for allowing us to form a network of casual friendships. But this is not the only aspect of honesty that benefits its possessor. Honesty also underwrites trust, and the ability to trust is necessary to form close relationships. Because these close relationships are important to our well-being, there is an egoistic reason to cultivate honesty. Honesty is a self-regarding virtue.⁵

⁵ Thanks to Doug Lewis for comments on an earlier version of this chapter.

Chapter 7

Character Egoism and Social Justice

I have argued, throughout, that the virtues are character traits that reliably benefit the trait's possessor. Part of the motivation for this view of virtue is the intuition that the catalog of virtues should be sensitive to context in a way that Aristotelian virtues are not. A diplomat should cultivate dispositions that allow her to move comfortably between cultures, while this trait is totally irrelevant for a farmer who has no interest in travel. And so on. The Aristotelian catalog, which is grounded in excellence of *kind*, doesn't accommodate the intuition that there are different slates of virtues for different people in different circumstances.

While the adaptability of the egoistic catalog is appealing over many different circumstances, we should worry if there are some circumstances—especially circumstances that commonly occur—that would cause a despicable set of traits to be viewed as virtues. If, for example, in circumstances of serious, widespread, or institutionalized *oppression* a repugnant set of character traits were reliably beneficial to the agent, then the theory I have sketched is in serious conflict with our

pre-theoretic intuitions. We admire people who fight for social justice *especially* when they themselves are willing to suffer in justice’s pursuit.

Theories in the Aristotelian tradition can accommodate our intuition that those who struggle against oppression are morally praiseworthy. Those who fight for justice in the face of oppression, displaying courage while maintaining kindness—that is, those who embody the Aristotelian virtues even in terrible circumstances—display a strength of character that is inspiring. Against the odds, they embody a fine instance of the human kind in a hostile environment.

For a virtue ethic grounded in well-being, the existence of oppressive circumstances present a two-fold worry: on the one hand, we might worry that this picture of the virtues will encourage oppressed people to adapt to their oppression, rather than to resist it. On the other hand, we might worry that this same advice to adapt will encourage members of the oppressor class to embrace the role, rather than to reform unjust social structures.

I think both worries can be answered. In this chapter, I offer a self-regarding account of the character traits that support social justice.

7.1 Oppressed groups and the pressure to adapt

7.1.1 Challenge

Consider the disposition to anger: a disposition that has secured for itself a position in the hall of fame of vices.

Gabriele Taylor, in *Deadly Vices*, distinguishes two forms of deadly anger. First, anger might take the form of irascibility. The irascible man is frequently—constantly, in the extreme case—angry, and he displays his anger. Irascible people “are given to bursts of temper quite out of proportion to the occasion; they lose

control of themselves and so may do much harm, both to themselves and to others” (Taylor 2006, 83-84).¹

A second variety of dangerous anger is bitterness. Bitter people are those “who suppress their animosity and keep up their anger for a long time. They have to digest it internally and so labour under the weight of resentment” (Taylor 2006, 85). This weight of resentment is what proves destructive to the self and thus, in Taylor’s view, makes embittered, silent anger a deadly vice. “The resentful expect to be undervalued and tend to find their expectations confirmed. Hence their resentment will become ever more firmly established” (Taylor 2006, 90). At least under normal circumstances, then, chronic anger is a vice because it harms the angry. It harms them so much that we call it a *deadly* vice. Though the characterization of deadly anger Taylor offers is embedded in Catholic tradition, it’s a characterization that could have come from a character egoist: chronic anger is a vice *because* it undermines our own well-being.

Lisa Tessman, in *Burdened Virtues*, agrees that a high level of anger leads to terrible consequences for the angry. However, she argues that under oppression, anger (among other character traits) can be a virtue *even though* it undermines the well-being of its possessor.

If tremendous anger is ultimately unhealthy or corrosive for its bearer, then the political resister with an angry disposition displays an example of what I have been calling a burdened virtue: a morally praiseworthy trait that is at the same time bad for its bearer, disconnected from its bearer’s well-being (Tessman 2005, 124).

¹ Taylor does not consider this form of anger to be a deadly vice in its own right. Rather, irascibility is a corollary vice of excessive pride. The excessively prideful person sees slights lurking in every shadow, and so is constantly angry that he is not receiving his due.

It is easy to find examples that lend intuitive force to Tessman's view. Consider Kunta Kinte in Alex Haley's *Roots*. Unlike Fiddler, a man born a slave, who seems able to build a pleasant life within those confines, Kunta is angry, unwilling or unable to embrace his identity as a slave. His anger is admirable, even inspiring, and yet it comes with a clear cost. First, it is *pleasanter* to be calm than angry—Fiddler, submissive and docile, sure *seems* happier than Kunta angry—and second, Kunta would have been tortured less had he submitted without resistance. Though we, as readers, understand Fiddler's submissiveness, we admire Kunta's rage—it seems the appropriate reaction to the injustice he is forced to endure. Kunta's anger appears to be both praiseworthy and bad for him.

In *Burdened Virtues*, Tessman's focus is less on personal struggles like Kunta Kinte's than on the political struggles of activists. Tessman's political resisters are outraged at oppression, and draw from their anger the energy to fight. But in the face of terrible oppression, it is unlikely that "raging political resisters can metabolize their anger; instead, I contend, they themselves suffer from the level of anger prescribed for them" (Tessman 2005, 124). Think of Frederick Douglass, an example to which I'll return shortly. Douglass made great strides in reforming the system of slavery into which he was born. But surely he himself suffered from the simmering (and occasionally explosive) anger that helped fuel his struggle.

Tessman's recent work follows on a large and growing body of feminist treatments of anger (Bell 2005, Spelman 1989, Narayan 1988, Frye 1983, Lorde 1984). Together, they make a compelling case for the oppression-resisting and change-fostering power of anger in circumstances of injustice. Tessman's book shows how a eudaimonistic virtue ethic like Aristotle's can be adapted to allow—or even require—virtuous agents to cultivate the character traits that serve the fight for justice even when these same character traits are bad for the agents themselves.

The challenge is this: if anger really is, as Tessman puts it, “disconnected from its bearer’s well-being,” then anger cannot ever be an *egoistic* virtue. But surely, in some circumstances, anger can be a virtue. Surely Kunta Kinte and Frederick Douglass are morally praiseworthy *in part because of* their anger at injustice. It is a mark against character egoism if it cannot account for the goodness of their anger.

7.1.2 Response

Let us take Frederick Douglass as an exemplar of the morally praiseworthy, chronically angry political resister. Douglass, famous for his demeanor of steady, controlled rage, describes the turning point in his life as a slave as a sudden blossoming of anger into physical violence. When attacked by Edward Covey, then his owner, Douglass for the first time fought back.²

[The battle with Covey] rekindled the few expiring embers of freedom, and revived within me a sense of my own manhood. It recalled the departed self-confidence, and inspired me again with a determination to be free. The gratification afforded by the triumph was a full compensation for whatever else might follow, even death itself. He only can understand the deep satisfaction which I experienced, who has himself repelled by force the bloody arm of slavery. I felt as I never felt before. It was a glorious resurrection, from the tomb of slavery, to the heaven of freedom. My long-crushed spirit rose, cowardice departed, bold defiance took its place; and I now resolved that, however

² Thanks for the example to Macalester Bell, who began a discussion of anger by considering the case of Frederick Douglass, via the following quote, in her “Anger, Virtue, and Oppression,” presented at the Minnesota International Conference in Ethics in 2007.

long I might remain a slave in form, the day had passed forever when I could be a slave in fact. I did not hesitate to let it be known of me, that the white man who expected to succeed in whipping, must also succeed in killing me (Douglass 2000, 74-75).

After this incident of explosive rage, Frederick Douglass cooled to a state of chronic (or settled, or characteristic) anger. Douglass surely suffered from his chronic anger, and yet his anger is part of what we admire in him.

How does Tessman account for the virtuousness of Douglass' chronic anger? The answer goes all the way back to Aristotle.

Aristotle identifies a virtue of appropriate anger, a character trait that urges its possessor to be angry "at the right things and with the right people, and further, as he ought, when he ought, and as long as he ought" (NE 1126a9-27). Tessman, summarizing *Rhetoric* Book II, writes that "anger, for Aristotle, is a feeling of pain at being unjustifiably harmed by another, especially if one is harmed by being slighted, that is, denied the respect that one deserves, and this pain is mixed with pleasure at the thought of revenge" (Tessman 2005, 120). This idea of respect is central to Aristotle's account of anger. Indeed, the appropriately angry man "is angered only by those who ought to show him respect; and he is angered only on such occasions when those who ought to show him respect do not do so" (Frank 1990, 272).

The connection between anger and self-respect is broadly acknowledged in feminist literature. In an overview of feminist views of anger, Macalester Bell writes, "[anger] may help women retain their self-respect. To accept meekly being wronged and not to respond with anger often signals weakness, whereas responding to wrongdoers with resentment or anger indicates that a person has at least a modicum of self-respect" (Bell 2005, 81).

Tessman quotes a series of feminists who make similar points. Marilyn Frye writes that, in responding to oppression with anger, one “claims that one is in certain ways and dimensions *respectable*. One makes claims upon respect” (Frye 1983, 90). Elizabeth Spelman writes:

To be angry at him is to make myself, at least on this occasion, his judge—to have, and to express, a standard against which I assess his conduct. If he is in other ways regarded as my superior, when I get angry at him I at least on that occasion am regarding him as no more and no less than my equal. So my anger is in such a case an act of insubordination: I am acting as if I have as much right to judge him as he assumes he has to judge me (Spelman 1989, 266).

The language of respect is also reflected in the passage from Frederick Douglass, quoted above. According to Douglass, the blossoming of violent anger, directed at Covey, “recalled [his] departed self-confidence,” allowed his “long-crushed spirit” to arise, and “revived within [him] a sense of [his] own manhood.” For Frederick Douglass, self-respect and anger at oppression are linked.

Imagine Frederick Douglass taking Covey’s assaults in stride. It seems that if he did, he would fall under Aristotle’s description of inirascibility, a vice.

Those who are not angry at the things they should be are thought to be fools, and so are those who are not angry in the right way, at the right time, or with the right persons; for such a man is thought not to feel things nor to be pained by them, and, since he does not get angry, he is thought unlikely to defend himself; and to endure being insulted and put up with insult to one’s friends is slavish (NE 1126a3-8).

Tessman inherits this analysis of anger from Aristotle, and adapts it to the sort of oppressive circumstances that seem to have been of little interest to him—circumstances like those of Frederick Douglass. Tessman writes:

People of color may be subjected to racist insults and degradation daily, and under such conditions, the “proper” level of anger for people of color (or their friends/political allies) becomes, relative to the anger appropriate to those who do not regularly encounter such insults, extreme. If one believes (correctly) in one’s own moral worth while others in the society—in dominant positions—do not believe in it, one will constantly find oneself “slighted” (to use Aristotle’s term); the frequent or unabating nature of this slighting... are conditions that make the right level of anger a tremendous level, the level of fury or rage (Tessman 2005, 123).

Frederick Douglass provides, I think, a model case of Tessman’s praiseworthy chronic anger. His is, at bottom, Aristotelian appropriate anger placed in circumstances that call for anger nearly constantly.

But is it the case that when we respond with admiration to the scene between Douglass and Covey, we are responding to Douglass’ *anger*? Suppose, instead of trying to elicit a response of meek compliance from Douglass, Covey had *tried* to elicit a response of anger. Had Covey goaded Douglass, trying to force him to lose control by losing his temper, I submit that we wouldn’t respond to anger with admiration. We would react with understanding, surely, but not admiration. This is because, were Covey *trying* to elicit anger, a response of anger wouldn’t have been *defiant*. This should make us suspect that it is the *defiance* of oppression that we admire, not the anger *per se*.

In fact, we have, in the case of Nelson Mandela, a relevant real-life example. While in prison in South Africa, Mandela's jailers tried to break his spirit by enraging him. He refused to become angry, and reacted to his jailers with firm, but quiet, dignity.³ Mandela's jailers deserve anger in the sense that they are appropriate targets for a whole range of negative reactive attitudes, outrage and anger among them. But, in the case of Mandela himself, his defiance of their assault on his dignity cannot be implemented through anger, since that is the very reaction they seek.

What we admire, in all these cases, is a level of self-respect that cannot be dominated, that responds to oppression with defiance. We admire a character that is unable to accept, or take up, an insult. Self-respect, when challenged by oppression, often, though not always, manifests itself as anger. When self-respect is instantiated as anger, it is what Aristotle labeled as appropriate anger. But the same self-respect, the same defiance of oppression, can be instantiated in different affects, such as pity or quiet fortitude.

If we classify anger as a virtue, we mistake the symptom for the cause. Profuse sweats under some circumstances—reading at the office, say, or sitting on the couch watching a movie—suggest ill-health. Profuse sweats under other

³ See, for example, Frontline's interview with Fikile Bam, who was imprisoned with Mandela for a decade:

No, I never saw him fly into a rage. I've seen him angry and I've seen him being very firm, and that's as far as it went. He never lost his cool, even when he was angry. In fact, I recall that he tended to be softer in his voice when he wanted to make a point, but firmer. This is how he treated warders, even warders who were nasty. He didn't shout at them ... I've never heard him swear at them, but he would just tell them and give them a very good lesson and a very good lecture in how he wanted to be treated and how he didn't want to be treated. How he expected decency and how he had been fighting for decency all his life and that he was not, now, going to abandon that particular struggle (Frontline 1999).

circumstances—digging post-holes on a hot day—suggest good health. Anger, too, is a symptom whose import changes with circumstance.

Aristotle and Tessman alike have mistaken the most common manifestation of an underlying virtue for the virtue itself. In the case of Frederick Douglass, we do not admire his *anger*, we admire his self-respect, which is appropriately manifested as anger. In the case of Nelson Mandela, we do not admire the quiet fortitude he shows his jailers, we admire his self-respect, which is appropriately manifested as quiet fortitude.

Let us return, for a moment, to Gabrielle Taylor’s characterization of irascibility—the form of deadly anger that manifests as “bursts of temper quite out of proportion to the occasion.”

The irascible man is very conscious of possible slights, and so his anger will be excessive also in respect of the range of occasions to which he responds with this emotion; excessive, therefore, in which he regards as an insult to be taken seriously. For him the world is full of people refusing to give him the respect he considers due to him, an attitude which in turn implies that he is much concerned with the importance of his own position (Taylor 2006, 84).

The slights that the irascible man exaggerates or gins up from air are real in the case of Frederick Douglass. First as a slave, and second as a black man in a society of racial slavery, Douglass was, nearly constantly, suffering assaults on his self-respect. The very social organization of the United States at the time was itself an unjustifiable assault on his self-respect. Unlike the irascible man, Douglass’ constant anger is an appropriate response to an unending series of real insults, each of which should elicit a response of anger.

Whereas chronic anger in irascible people is a symptom of excessive pride, or an out-sized concern with commanding deference from others, chronic anger in oppressed people is often a symptom of secure and stable self-respect.

Self-respect is something close to a necessary condition of well-being. Without, as Rawls puts it, a “sense of his own value, his secure conviction that his conception of the good, his plan of life, is worth carrying out” (Rawls 1971, 440), a person can’t hope to pursue any projects or relationships that are specially important to himself. For members of an oppressed group, *adapting* to the circumstances of oppression would entail sacrificing self-respect. Because self-respect is unambiguously an egoistic virtue—and a central one—we should expect that character egoist agents will resist oppression in most circumstances.

In short, from the perspective of well-being, adaptation to circumstances of oppression is, in most cases, a losing bargain. The angry affect of people like Frederick Douglass, though it comes at a clear cost, is the best option in a tragic situation. Preserving his self-respect is centrally important to his ability to live a life that is worthwhile to him, and if that comes at the cost of a disposition of settled anger, it is worth it. Tragic, but worth it.

7.2 Social justice as a self-regarding virtue

7.2.1 Challenge

Skepticism about the first-personal benefits of justice gets full-throated expression by Glaucon in the *Republic*. In many ways, Glaucon’s skepticism about justice parallels the sensible knave’s skepticism about honesty.⁴ Glaucon asks us to

⁴ The central difference between the vices of dishonesty and injustice is that dishonesty analytically requires deception, while injustice does not. Thus, my response to the sensible knave,

consider a man who has secured a reputation as impeccably just, who privately denies the rights of others, and secretly exploits them when he can. Isn't this man better off than the *actually* just man, who is constrained by his consideration for the rights of others?

Philippa Foot, in "Moral Beliefs," offers one line of response to Glaucon's skepticism about justice. If we are to undertake a campaign of injustice, she says, we are faced with two options. On the one hand, we can carry out our campaign in the open. But to be openly unjust means we will have to fear similar treatment at the hands of others; if we treat other people brutally, we can expect brutal treatment in return. On the other hand, if we follow Glaucon's advice and keep our injustice secret, we must expend huge amounts of effort in cloaking ourselves. Anyone at any time could discover our true nature, and blow our cover. In order to avoid such discovery, "the price of vigilance would be colossal" (Foot 1967, 100). Thus, says Foot, because both options are terrible, it is prudent to avoid them by being just.

In "How Injustice Pays," Bernard Boxill argues, contra Foot, that Glaucon's strategy of profiting from injustice is much easier to carry out than Glaucon himself seems to believe. This is because deception need not be part of the strategy. Rather, we can reap the benefits of justice and injustice simultaneously in the following way. First, find a group of people with an obvious—which is to say, recognizable at a glance—difference. Second, agree to be genuinely just with people who look like you, and openly unjust toward members of the different group.

This strategy of limiting the scope of justice, so as to profit from the exploitation of people who fall outside its boundaries, is, of course, the form that injustice most commonly takes in the real world. But once we recognize this, Philippa Foot's

which focuses on deception's tendency to degrade our ability to trust others, isn't available as a response to Glaucon.

response to Glaucon falls flat. Injustice can be carried out in the open, and so long as the unjust person “persecutes only one other, or only certain well-defined others, it is not clear that all others must fear him and accordingly that he must fear all others. Perhaps only his victims need fear him and he need fear only his victims” (Boxill 1980, 360).

Boxill’s claim is that injustice, especially when institutionalized, can be compartmentalized. An unjust relation to one person or group does not require an unjust relation to every person or group.

There is a sense in which Boxill is obviously, unquestionably right. Black slavery in the Antebellum South was never accompanied by white slavery. White owners of black slaves did not find their white friends withdrawing, out of fear that they, too, would be enslaved. This is because white slave owners effectively compartmentalized their injustice: considerations of justice bound interactions with whites, while considerations of expediency alone bound interactions with blacks.

Boxill’s contention is that an oppressive arrangement, like slavery, can be stable, and if stable, can be beneficial to the oppressors. An arrangement in which a plantation owner benefits from slave labor, without himself incurring any risk of being enslaved, is a profitable one for the plantation owner. This is one way injustice can pay.

If we consider the costs and benefits of reform from the perspective of a member of the oppressor class, it quickly becomes clear that sacrificing the material benefits that follow on institutionalized injustice is only one aspect of the full cost of reform. After all, the agreement about the boundaries of justice is mutual among members of the oppressor class, and works to their material benefit. To fight to overturn the oppressive agreement risks alienating the resistor from her peers. In fighting

to end slavery, the Southern abolitionist, for example, risks not only her financial security, which rests on the back of slave labor, but also her relationships with her friends and family.

These sorts of cases—cases of people who find themselves members of an oppressor class—present a problem for character egoism. If the virtues are those traits of character that maximize the probability that their possessor will have successful relationships and projects, then what are the virtues of the oppressors? We might worry, given the material benefits of oppression and the costs of reform, that character traits that perpetuate oppression will be *virtues* for the oppressor class. But this can't be right. Surely the most admirable people are those who fight to end injustice, *despite* the costs to themselves.

The challenge is this: can it possibly be the case that the character traits that urge reformers to abandon material benefits and alienate their peer group actually leave them better off, in the end, than the character traits that would allow them to enjoy, peaceably, the fruits of oppression? And if not, then doesn't character egoism classify as vices the very character traits that tend to promote the reform of unjust institutions?

7.2.2 Response

Those who benefit from oppressive social structures are oriented toward those structures in different ways. I see at least three clearly different orientations the beneficiaries of oppression can have toward the oppression itself. They might be directly involved in establishing or maintaining the oppressive arrangement. Or they might be so well insulated from the business of oppression that they don't even know it's happening, though they still benefit from it. Or they might be aware that they benefit from the oppressive arrangement, and choose to avert

their gaze, doing their best to put out of their minds the suffering of oppressed groups, at least in part in order to continue reaping oppression's benefits. Different issues arise from each of these orientations.

Active supporters

Oppressive social structures need active supporters if they are to persist. Many people recognize that they are in a dominant position in society, and fight to maintain it.⁵ This active support of oppressive social structures can take many forms, from working as an overseer of slaves to protesting in support of gay marriage bans, from publishing anti-suffrage pamphlets to voting for racist politicians. Let us take up this question first: is character egoism compatible with active support of unjust social structures?

To begin an answer, we should note, first, what seems to be a simple fact of human psychology: compartmentalizing injustice takes effort. Boxill suggests that the difficulty of compartmentalization explains why real-world exploitation so often falls along lines of race. Quoting first Rousseau, then Martin Delany, he writes:

People are far more likely to render justice to those they readily identify with, and far less likely to suffer for the injustice they do to those they do not readily identify with. For pity is the “innate abhorrence to see beings suffer that resemble” one and is more potent when the one who “beholds the distress identifies himself with the animal that labors

⁵ I'm simplifying, here, for convenience. There are many examples of groups on the downside of oppression who fight to maintain the oppressive arrangement. Women were well-represented in the anti-suffrage movement, for example. Though these cases may be psychologically complicated, I don't think they change what I have to say, here. The point is that there are some people who recognize, and actively support, oppressive social structures.

under it.” Thus, endeavoring to explain how Europeans could enslave Africans with such apparent equanimity, Martin Delany, the first important black nationalist, wrote that oppressors always “select as the objects of proscription, those who differed as much as possible from themselves” for this “ensures the greater success of the project because it elicits less interest on the part of the oppressing class” (Boxill 1980, 364-365).

Though sympathy operates at its strongest on people that closely resemble us, its scope is far, far wider than that. We have sympathetic responses to animals, for example. And though our sympathetic responses are strongest with large mammals like polar bears—animals relatively close to us in the taxonomic tree—we still do have sympathetic responses to birds with broken wings and lizards cowering in apparent fear.

While it is true that it is easier to withhold sympathy from people with obvious superficial differences from us, withholding sympathy still takes effort. What is the long-term effect of withholding sympathy from people, or other creatures, that naturally elicit sympathetic responses? One likely outcome is a numbing or deadening of sympathy. We certainly expect to see this deadening with respect to the subject from whom sympathy is withheld. We would expect a meat packer, after a few months sloshing around in the abattoir, to cultivate indifference to the happiness or suffering of pigs and cows. Likewise, we would expect a slave overseer, after a time spent treating black people as property, to cultivate indifference to the happiness or suffering of black people.

This deadening of sympathy is a psychological *defense mechanism*, though. The reason the slaughterhouse worker struggles to become indifferent to the suffering of pigs is because, were he to let his sympathy operate unhindered, his life

would be intolerable.

When faced with a situation in which I suffer in sympathy with members of an oppressed class, I have at least two avenues available to me to end my own pain. On the one hand, I could end the suffering, so there is no longer any pain for my sympathy to respond to. On the other hand, I could block the normal functioning of my sympathy, so that it doesn't convey the suffering, though the suffering persists.

The first option—ending the suffering—requires the ending of the oppressive arrangement itself. This has associated costs, discussed above, both material and interpersonal.

But we might suspect that the second option also has a dear price: generalized damage to our ability to sympathize. If I deaden my sympathy toward one group of people, do I risk broader damage to my sympathetic sensitivity? Does the psychological mechanism of sympathy respect the superficial distinctions on which we base our oppressive arrangements?

The question, here, concerns how effectively our *sympathy* (not our acts of injustice) can be compartmentalized. If it is really possible that white oppressors can be indifferent to the suffering of black men and women without that harming the operation of their sympathy more generally, then it may well be the case that injustice can pay, all things considered. But there are reasons to believe sympathy toward one group cannot be deadened without risking serious harm to the psychological mechanism of sympathy more generally.

In *The Jungle*, his novel set in the early days of the factory model of animal slaughter, Upton Sinclair writes that “men who have to crack the heads of animals all day seem to get into the habit, and to practice on their friends, and even on their families, between times” (Sinclair 1906, 19). There is some evidence that

Sinclair is right.

Increases in crime—property crime, violent crime, domestic violence—in slaughterhouse towns have been observed throughout North America. In her dissertation exploring the causes of the observed increase, Amy Fitzgerald argues that none of the existing explanations can account for the facts. Previously, social scientists have suggested that increases in crime rates could be due to increases in the immigrant population (from which slaughterhouses usually draw), or more generalized population booms, or increased unemployment from the high rate of turnover associated with slaughterhouses. According to Fitzgerald, none of these explanations accounts for the observations. She suggests a new explanation: “the unique work of killing and dismembering animals undertaken in slaughterhouses might result in negative spill-over effects into the larger community” (Fitzgerald 2006, 195).

The statistical tools that Fitzgerald employs reveal the shortcomings of the existing explanations, but cannot make much progress in establishing her own psychological explanation. To this end, Fitzgerald quotes from Gail Eisnitz’s interviews with former slaughterhouse workers. These workers seem to confirm that they struggled not to fall victim to the psychological mechanism described by Upton Sinclair. One stick-pit worker—that is, someone responsible for slashing the throats of the pigs—writes:

The worst thing, worse than the physical danger, is the emotional toll. If you work in that stick pit for any period of time, you develop an attitude that lets you kill things but doesn’t let you care.

Another writes:

My attitude was, it’s only an animal. Kill it. Sometimes I looked at

people that way, too. I've had ideas of hanging my foreman upside down on the line and sticking him. I remember going into the office and telling the personnel man that I have no problem pulling the trigger on a person—if you get in my face I'll blow you away (Eisnitz 1997, 87).

When it comes to slaughterhouse work, the prospects for compartmentalization don't look good. A mountain of anecdotal evidence, and Fitzgerald's social psychological research both indicate that the daily routine of acting brutally toward pigs and cows affects workers in such a way that they often have to struggle not to act brutally toward people—even their families and friends.

A comparable study cannot be done—thankfully!—on communities in which slaveholding is newly introduced. If Rousseau and Delaney and Boxill and Hume are right, though, we would expect the sympathetic response to human suffering to be even more difficult to bear than the slaughterhouse workers' response to animal suffering. Indeed, this effect seems borne out in accounts of slavery written at the time.

Frederick Douglass sketches the characters of several slave owners and overseers. In most every case, the overseers and owners well-suited to their brutal work are ill-suited to more benign circumstances. Douglass describes his first owner, Captain Anthony, as “a cruel man, hardened by a long life of slaveholding” (Douglass 2000, 20). Douglass' passing comment suggests that Captain Anthony's hardness and cruelty are general character traits, not localized hardness-to-slaves and cruelty-to-slaves.

Douglass offers more detail about the character of Austin Gore, an overseer “possessing, in an eminent degree, all those traits of character indispensable to

what is called a first-rate overseer.” Says Douglass, “Mr. Gore was proud, ambitious, and persevering. He was artful, cruel, and obdurate. He was just the man for such a place, and it was just the place for such a man” (Douglass 2000, 33). This man, so well-suited to managing slaves, “was a grave man, and, though a young man, he indulged in no jokes, said no funny words, seldom smiled” (Douglass 2000, 34). In fact, he is isolated by his coldness, as when the community recoils from his murder of an unarmed slave. “A thrill of horror flashed through every soul upon the plantation, excepting Mr. Gore. He alone seemed cool and collected” (Douglass 2000, 35).

Douglass is surely right that Austin Gore is suited for slave overseeing, and his fitness to this role is borne out in his success. But despite his success, Gore is a pathetic figure—someone few people on earth would be willing to trade places with. His successes in torture and intimidation come at the cost of relationships.

If our sympathetic responses don’t respect the compartmentalization we employ to make injustice pay, the choice faced by someone born into an oppressive arrangement looks a little different than it did at the outset. On the one hand, we can avoid the pain of sympathetic suffering by reforming the system to eliminate the oppression. On the other, we can deaden our sympathy toward those suffering, which risks deadening our sympathy more broadly.

If my own happiness or sadness has no effect on my companion’s state of mind, there’s little hope that companion could be a friend. This is generally true of most forms of interpersonal relationships: if two people do not or cannot sympathize with each other, they have few satisfying forms of relationship available to them. Because sympathetic sensitivity is a prerequisite for just about any personal relationship we could want, it is a good candidate for a core virtue.

Literature is loaded with characters whose indifference to suffering serves them

well in some narrow way, but harms them when we take the wider view. To return, again, to Ebenezer Scrooge: his indifference to poverty's depredations allows him to make a killing on the free market, but cuts him off from the human interactions that he eventually judges to be more important.

In these extreme cases—the people who work at the leading edge of oppression's wedge, the pig stickers, the slave overseers, and so on—the case for justice as a self-regarding virtue is strong. We all do better in our own lives if we don't directly cause the suffering of others. Character egoism, which encourages adaptation to circumstances with an eye toward well-being, will *not* encourage adaptation to the sorts of life roles that directly support unjust social institutions and arrangements, because such adaptation deforms our characters in a way that devastates our prospects for well-being.⁶

On why the preceding is not an example of a right answer for the wrong reason

Kant's view of the wrongness of animal cruelty includes two claims: 1) mistreating animals makes us callous. And 2) callousness increases the risk that we'll harm *rational agents* (that is, people) and this is what makes it wrong to mistreat animals.⁷ The first claim, though it remains controversial, is perfectly plausible

⁶ In the end, my response to the enthusiastic slave-owner will probably, frustratingly, share a feature with my response to the sensible knave: it will be unconvincing to the very people who most need to be convinced (see section 6.6). After all, owning slaves has clear material advantages. The harm to well-being that follows on dispositions of injustice, cruelty, and benumbed sympathy is much harder to quantify. I say "better to leave your sympathetic sensitivity intact than to grow wealthy on a tobacco plantation." Covey, Frederick Douglass' torturer, will likely respond that he prefers to keep the money. I suspect there is something tragically wrong about the guy, in much the same way I suspect there's something tragically wrong with the knave, who thinks gold is more important than friends. But until Covey has had the experience of the kinds of relationships that are currently closed off to him due to his cruelty and callousness, it is unlikely that my response will get a grip on him.

⁷ See *The Metaphysics of Morals* 6:443:

and always has been. It's the second claim that is embarrassing to some defenders of Kant. He's just totally wrong about what makes torturing animals wrong. The problem with causing needless suffering to animals is that it causes needless suffering *to animals*. The problem with cruelty to animals is that it's cruel *to animals*.

Does my character-egoist account of why injustice is wrong include a similarly embarrassing mistake? It certainly would if the claim were that people should not buy slaves (for example) because they might accidentally degrade their family relationships thereby. The problem with slavery is that it is cruel and unjust *to the people being enslaved*.

I do not mean to suggest that the problem with slavery is that it undermines relationships with your friends and family. Slavery is bad because it is unjust and cruel. My point is that the deformative effects of these vices are easiest to see when we consider those vices imported into relationships we care about. We can see why injustice and cruelty are bad *for us* by looking at their effects on relationships we especially care about.

The difference between a self-regarding account of justice and Kant's account of duties to animals is in which layer in the strata of norms the relevant norms figure in. We generally read Kant as arguing that respecting rationality is the top layer of justification. Question: "Why didn't you kick that dog just now?"

Ridiculous answer: "In order to reduce the odds that I'll kick a rational agent in the future."

Violent and cruel treatment of animals is... intimately opposed to a human being's duty to himself...; for it dulls his shared feeling of their suffering and so weakens and gradually uproots a natural disposition that is very serviceable to morality in one's relations with other people.

The self-regardingness of ostensibly other-regarding virtues like justice and kindness comes in, for a character egoist, at a different level of justification. “Why didn’t you kick that dog just now?”

“Because he’s a dog doing doggy things. What kind of monster kicks a dog for doing doggy thing?”

“Why not just become cruel, and take pleasure in kicking dogs?”

“Because that sounds like a miserable way to live.”

Similarly, “Why don’t you own any slaves?”

“Because that would be a disgustingly unjust and cruel way to treat other people.”

“Why not cultivate an attitude of disdain for considerations of justice, so that you can then merrily buy slaves and make them do your work?”

“Because that sounds like a miserable way to live.”

I say again: adapting to unjust situations deforms our characters in ways that harm our prospects for well-being. Thus, agents who accept character egoism have reason to cultivate character traits that foster social justice, and thus have reason to refuse to maintain or promote unjust social structures.

The honestly ignorant

The worry about adaptation to oppressive structures still persists, in attenuated form: it is well and good that character egoism counsels against our being active oppressors. But many people in oppressive societies live much farther from oppression’s front lines than do Ebenezer Scrooge and Austin Gore. For people who are well-insulated from exposure to members of oppressed groups, blissful ignorance is a real possibility. The psychological costs of supporting oppression that I have highlighted, in the case of slave overseers and pig stickers, depend on

exposure to the suffering of the oppressed. We cannot respond sympathetically to people we never see.

One might suppose that this is a common state of affairs today with respect to factory-farmed animals. While ignorant of the suffering of these animals, it's easy to enjoy cheap meat. But of those who get a glimpse of the conditions on factory farms, many choose to give up meat, checking out, at least for a while, of the factory farming system.

In the *Metaphysics of Morals*, Kant recommends that we tour poverty-stricken neighborhoods and debtors' prisons, so that our natural sympathy will help drive us to fulfill our duty of beneficence (Kant 1996, 6:456). It appears that character egoism lacks the resources to make such a recommendation. Because it is difficult to see how, in normal cases, a broader knowledge of oppression and injustice outside my normal sphere of experience can reliably foster my own well-being, it is difficult to see how any self-regarding disposition could urge us to tour debtors' prisons and high-density feed lots.

To be clear: there are, as I have argued, good self-regarding reasons to be kind and sympathetically sensitive to others. When a kind and sensitive person sees someone suffering due to oppression, we would expect that she'd be bothered by it. So character egoism classifies as virtues those traits of character that cause us to recoil when we are exposed to injustice.

What character egoism appears to lack is a virtue something like a disposition to seek out injustice. Character egoism doesn't appear to condemn the complacent ignorance of a kind, sympathetically sensitive person who doesn't realize she is benefiting from the operation of unjust structures she never sees.

However, there is good reason to think that the disposition to seek out injustice *isn't a virtue* that could survive reflective scrutiny. That is, I think that this

character trait, unlike honesty and kindness, is not a piece of moral data that a theory of the virtues must justify. I think this for two reasons.

First, imagine a white family living on a self-sustaining farm in the slave-holding south. 1) They have no slaves of their own. 2) They know, in the theoretical sense, that nearby communities hold slaves. 3) They have the full slate of core virtues like honesty, kindness, self-respect, and sympathetic sensitivity, and as such, if they were exposed to slavery in action, they would be horrified, and would be motivated to work in earnest for its abolition. But, 4) they have no such exposure, and though they occasionally trade with people in towns who do own slaves, they have met neither slaves nor slave-holders, they never see the brutal side of slavery, and don't go seeking it out. In this situation, calling these people *virtuous* people doesn't seem nearly as controversial as it would be were I to sketch a parallel example in which a person had all of the core virtues except, say, kindness. This should suggest that the disposition to root out injustice is not a core virtue.

Second, to build the disposition to root out injustice into our pre-theoretic portrait of virtuous people opens the door to the same sort of objections that have dogged action-focused theories like utilitarianism; it risks introducing *demandingness* into virtue theory.

Demandingness is a problem of whole moral theories, not of particular requirements of a theory. There are no one-time requirements that can be ruled out, *a priori*, as too demanding. Peter Unger convincingly makes this case in *Living High and Letting Die*. He imagines difficult actions—terribly difficult, costly actions—and then constructs thought experiments in which the difficult actions are clear moral requirements. For example, he tells us a story about Bob, who recently sunk his entire retirement account into the purchase of a Bugatti. Through the sort

of coincidences that predictably crop up in philosophical thought experiments, Bob finds himself in a situation in which he has to decide between wrecking his Bugatti and allowing a child to die. He chooses the car over the child. According to Unger, “when confronting this engaging example, everyone responds that Bob’s conduct was monstrous” (Unger 1996, 136). Surely this is right. Even when it means sacrificing his future financial security and thus, almost certainly, his future well-being, we think it would be monstrous of Bob to let the child die. The general upshot of Unger’s examples is that in any particular case, morality can be extremely demanding without being *problematically* demanding.

Nor is a theory demanding (or, at least, not necessarily so) if it *consistently* hands down obligations that are very difficult to discharge. There are at least two reasons an agent could find a given morality consistently difficult to live up to. It could be that she is a bad person, for whom moral standards are difficult to meet, even when they aren’t difficult for others. Or, the standards of morality could be difficult for anyone—even good people—to meet.

It doesn’t make sense to accuse a theory of demandingness in the first sort of case. Imagine a case in which a mean-spirited person repeatedly finds herself in situations in which morality requires her to be kind. The fact that she will repeatedly struggle (or fail) is no reason to think her morality should not require kindness of her.

Similarly in the case of a naturally kind person. There is at least one straightforward reason why a kind moral agent might still have to struggle to be kind: it might be that the world is simply such an imperfect place that the amount of kindness required to live morally in it would be tiring for *anyone*. In the case of the good moral agent, noticing that her theory is hard for her to live up to is an indictment of the world in which she lives. It is not by itself enough to show that

her theory is problematically demanding.⁸

The problem of demandingness, then, is not a matter of the *strength* required to live a theory. When philosophers like Williams (1973), Stocker (1997), and Wolf (1982) write about the effects of demanding theories on moral agents, they are writing about something else. These critiques do not target moralities that are *difficult* to live, but rather moralities that, *were they lived*, would render their agents lousy people. They are concerned with moralities that lead to the degradation or destruction of their agents' *characters*.

For Williams, the problem with utilitarianism (e.g.) is not that it can be difficult to live up to, but that if one were to live up to it, one would have to sacrifice all of one's ground projects. If we imagine someone with no ground projects—someone who has no relationships they cultivate, no talents they refine, no overarching goals they seek to achieve—we imagine someone who is boring, mechanical, two-dimensional, lacking integrity (Williams 1981, Williams and Smart 1973).

In “Moral Saints,” Susan Wolf pursues a similar line. She notes the activities, like reading Victorian novels, that can be crowded out by a demanding moral theory, and concludes that

⁸ Not everyone agrees that demandingness is not about demands on personal strength. Among virtue ethicists, Christine Swanton is an important exception.

Morality is demanding if it is demanding on the strength of the agent in three respects: she needs to be very strong to fully cultivate the virtues required in such a morality; she needs to be very strong to bear well the costs to her in exercising these virtues; and, given that she herself has not yet cultivated the virtues appropriate to the morality, she needs to be very strong to motivate herself to perform those acts that the virtues require (Swanton 2003, 203).

My disagreement with Swanton over whether difficulty in living a theory constitutes a problematic form of demandingness is fairly fundamental. I won't argue for my own view, but rather highlight that Swanton's worry is distinct from the Wolf/Williams/Stocker group of worries that I take to be characteristic of the virtue ethical demandingness objection to (especially) utilitarianism.

although no one of the interests or tastes in the category containing [Victorian novel-reading, e.g.] could be claimed to be a necessary element in a life well lived, a life in which none of these possible aspects of character is developed may well seem to be a life strangely barren (Wolf 1982, 421).

Similarly, for Stocker, the problem with utilitarianism (e.g.) is that it requires that its agents have a gap between the things they value and the reasons that move them to act. When we imagine an agent who can never be motivated by their values, we imagine an agent who is cold, passive, or in Stocker's terms, morally schizophrenic (Stocker 1997).

The concern of all three authors is, at bottom, the same: it has to be a mark against a moral theory if it turns its agents into the sort of people we would never want to be, or befriend, or be around. It has to be a mark against a theory if it turns its agents into bad people.

If this is the right description of the problem of demandingness, then it seems that virtue ethics can't be demanding. One of the core claims of virtue ethics is that character evaluation is prior to action evaluation; the moral quality of an action is understood in terms of the aspects of character it expresses (or fails to express). Virtue ethics' main concern is that its agents cultivate virtues and refine their characters, while the demandingness critique takes to task moral theories that lead to the degradation of their agents' characters in the pursuit of some other value.

Returning, at last, to the question of whether or not a moral theory ought to require its agents cultivate dispositions that urge the rooting out of injustice: one reason to be suspicious of the claim that morality *should* require us to root out oppression is that it seems to introduce the threat of demandingness into virtue

ethics.

It's the "rooting out" part that's the problem. Being pained by injustice is good. Being disposed to seek out opportunities to be pained is maybe not. Similarly, being moved by need is good. Being disposed to seeking out opportunities to be moved by need maybe not. Any of these seeking-out dispositions risks the destruction of good characters in much the same way Utilitarianism does.

Given that the world is a big place, full of oppressions large and small, open and hidden, there is a real worry that a drive to find and correct oppression will dominate our time and attention. If, to be judged good, we must root out oppression—an endless process, surely—it is hard to see how we can have time for friends or family, or for any projects other than the eradication of oppression.

Exposure to serious oppression should—and does, if we are decent people—seize our attention. Our sympathy for others makes it difficult or impossible to be happy when we are around others who are suffering. This can be a good thing if our exposure to oppression is intermittent—it helps urge us to reform institutions and eliminate the oppression. But if the oppression and attendant suffering is constant, it risks either dominating our attention or deadening our sympathy.

To incorporate the slogan that "no one is free while others are oppressed" not just cognitively, but affectively—to feel discomfort or pain at the knowledge that somewhere nearby, people are probably living under oppression—risks turning our lives into a single pursuit in much the same way utility-maximization does.

Some utilitarians—notably Peter Singer—bite the demandingness bullet and insist that, because the world is in such a terrible state, there is currently no moral room in it for Wolf's well-rounded people. It might be the case that some virtue ethicists prefer to bite the analogous bullet and accept that so long as serious oppression exists in the world, the only way to be moral is to relentlessly suss out

oppression. To take this route would sacrifice much of what is appealing about the virtue approach to morality.

In sum: character egoism doesn't classify as a vice the complacent ignorance of people who benefit from oppression without knowing it. But it isn't at all clear that a virtue theory *should* classify this kind of ignorance as a vice, because a drive to discover oppressions that are presently unknown risks sacrificing a whole character in pursuit of the lone value of justice.

The willfully ignorant

The third class of people who benefit from oppressive structures is the most problematic for character egoism. These are people who know they benefit from unjust arrangements, but choose to avert their eyes from the injustice, to put it out of their mind. This character was efficiently sketched (in an entirely different context and to entirely different ends) by Mencius:

The attitude of a gentleman towards animals is this: once having seen them alive, he cannot bear to see them die, and once having heard their cry, he cannot bear to eat their flesh. That is why the gentleman keeps his distance from the kitchen (Mencius 2004, 10).

It is easy enough to imagine circumstances of oppression or injustice in which character egoism would council us to take the path of Mencius's gentleman. This is especially true in situations in which the unjust structures are entrenched, and there's little possibility of short-term change. Imagine an otherwise decent person, embedded in an unjust culture, is faced with a choice of how to be. On the one hand, she could face up to the fact that she's benefiting from an unjust culture, be appalled by that fact, and struggle in vain for reform. On the other hand,

she could avert her gaze from the injustice, put it out of her mind, and live an otherwise decent life while accepting the unjust benefits of her culture. I do not see how to escape the conclusion that it would be better *for her* to avert her gaze—to stay away from the kitchen. And thus, I do not see how to escape the conclusion that the ability to avert her gaze is, by the lights of character egoism, a virtue.

Though I find this conclusion unpalatable, and would prefer to avoid it, I would like to close with a few words about why I do not take this to be a knock-down objection to character egoism.

First, note how circumscribed this issue is, relative to the broad skepticism about the first personal benefits of justice with which we began. And note that oppressive social structures present a problem for most theories. Aristotelian virtue ethics needs to be *adapted* for circumstances of oppression. (This is the topic of Lisa Tessman's book.) In short, it isn't clear what attitude we should adopt toward a theory of the virtues that works very well for just circumstances, but struggles in unjust circumstances.

Second, some help may lie in the (under-discussed) division of labor between moral and political philosophy. After all, the problem of egoistic virtues under circumstances of oppression is a problem of local maxima. It very well might be better for those placed in oppressive circumstances to develop the ability to put the suffering of the oppressed out of their minds, particularly if they are powerless to effect change. However, they could do better still if society were arranged in a non-oppressive way.

I think it isn't outrageous to believe that the role of moral norms should be to help agents achieve the highest maximum available to them, while the role of *political* norms is to make the highest of the maxima available to the most people.

That is, it is the responsibility of the *polity* to reform unjust structures within it. It misplaces the responsibility for social justice to expect individuals, who are themselves distant from the grind of oppression, and are themselves powerless to effect change, to undertake a futile effort.

This is not an answer to the problem posed by the willfully ignorant. It is, rather, a gesture in the direction of future work.⁹

⁹ Thanks to Valerie Tiberius, Matt Frank, and Mike Rohde for discussion about this material. Thanks to Steve Nelson for the Kant quote on duties to animals. Thanks to Jason Swartwood for the Mencius quote about gentlemen and their kitchens.

Chapter 8

Conclusion

I have argued that the virtues are best understood as those traits of character that reliably promote success in the projects and relationships that are (or ought to be, for her own sake) important to their possessor. This theory of the virtues as instrumentally connected to the well-being of their possessor is the core component of a character-egoist virtue ethic.

Work remains before character egoism can be counted a fully developed alternative to neo-Aristotelian virtue ethics. A few of the projects I have yet to take up:

A criterion of right action. Neo-Aristotelian theories generally derive a criterion of right action from a theory of good character: the right action is the action that would be undertaken by a similarly situated, fully virtuous person. Because character egoism accepts the possibility that two people could both be fully virtuous while each possessing a (somewhat) different slate of virtues, it's unclear whether the neo-Aristotelian approach can be adapted for character egoism.

The structure of the catalog of virtues. A consequence of character egoism is that all virtues are not equally important. To capture this, I have occasionally

used the language of “core” and “fringe” virtues. Roughly, core virtues are the sorts of traits that almost anyone will need to have if they are to do well in any project or relationship. Self-respect is the clearest example of a core virtue—it’s something close to a necessary condition of well-being for nearly everyone. Fringe virtues are traits that might be important for some people, or for some classes of project or relationship, but probably not for everyone. Wittiness, for example, might be critically important for some people, but many others get by just fine without it.

I’ve not yet begun to sort out the complications that follow on accepting a structured catalog of virtues. How closely does the set of core egoistic virtues match Aristotle’s catalog? Fringe virtues can be in tension with one another. How, then do we tell fringe virtues from minor vices? Is it possible that fringe virtues could conflict with core virtues?

Compatibility with evidence from psychology. My defenses of justice and honesty as self-regarding virtues turn on the premise that we are bad at compartmentalizing our dispositions and sympathies: it’s difficult for human beings to be dispositionally honest to members of an in-group and at the same time to be dispositionally dishonest to members of an out-group; it’s difficult for human beings to withhold sympathy from selected suffering people (or, for that matter, animals) while remaining sympathetically sensitive to everyone else. These are empirical claims, and although there is some research suggesting they are true, the evidence is, as yet, limited and equivocal.

This is the work that remains. I have tried to show, through the course of this dissertation, that this is work worth doing. I have argued that long-standing objections preempting the consideration of character egoism are not nearly as formidable as they are usually taken to be. Character egoism does not inherit

the problems of ethical egoism. It does not require us to abandon our reflectively endorsed intuitions about which traits of character are morally good, and which are bad. It has a plausible story to tell about how putatively other-regarding virtues like beneficence, honesty, and justice, do, indeed, benefit their possessor.

Beyond its innocence of the deficiencies commonly attributed to it, character egoism offers this signal advantage: it makes clear the content and the mechanism of virtue's reward and recompense. The reward of virtue is healthy relationships and successful projects. We could do worse.

References

- Annas, J.: 2003, Being virtuous and doing the right thing, *Proceedings and Addresses of the APA*.
- Annas, J.: 2008, Virtue ethics and the charge of egoism, in P. Bloomfield (ed.), *Morality and Self-Interest*, OUP, pp. 205–223.
- Appiah, K. A.: 2008, *Experiments in Ethics*, Harvard University Press.
- Argyle, M.: 2001, *The Psychology of Happiness*, 2nd edn, Routledge.
- Aristotle, I. t.: 2000, *Nicomachean Ethics*, 2nd edn, Hackett Publishing Company.
- Baier, A.: 1979, Good men's women: Hume on chastity and trust, *Hume Studies* **5**(1), 1–19.
- Baron, M.: 1982, Hume's noble lie: An account of his artificial virtues, *Canadian Journal of Philosophy* **12**.
- Bell, M.: 2005, A woman's scorn: Toward a feminist defense of contempt as a moral emotion, *Hypatia* **20**(4), 80–93.
- Boxill, B.: 1980, How injustice pays, *Philosophy and Public Affairs* **9**(4), 359–371.

- Calder, T.: 2007, Against consequentialist theories of virtue and vice, *Utilitas* **19**(2), 201–219.
- Cicero, G. t.: 1971, *On the Good Life*, Penguin Classics, chapter Laelius: On Friendship, pp. 172–227.
- Csikszentmihalyi, M.: 1990, *Flow: The Psychology of Optimal Experience*, Harper and Row.
- Darley, J. and Batson, C. D.: 1973, "from jerusalem to jerich": A study of situational and dispositional variables in helping behavior, *Journal of Personality and Social Psychology* **27**, 100–108.
- Darwall, S.: 1977, Two kinds of respect, *Ethics* **88**, 36–49.
- Diener, E. and Seligman, M.: 2002, Very happy people, *Psychological Science* **13**, 81–84.
- Dillon, R.: 1997, Self-respect: Moral, emotional, political, *Ethics* **107**(2), 226–249.
- Doris, J. M.: 2002, *Lack of Character: Personality and Moral Behavior*, Cambridge University Press.
- Douglass, F.: 2000, *Narrative of the Life of Frederick Douglass, an American Slave*, Modern Library.
- Driver, J.: 1999, Modesty and ignorance, *Ethics* **109**, 827–834.
- Driver, J.: 2001, *Uneasy Virtue*, Cambridge Studies in Philosophy, Cambridge University Press.
- Dworkin, R.: 2011, *Justice for Hedgehogs*, Harvard University Press.

- Eisnitz, G.: 1997, *Slaughterhouse: The shocking story of greed, neglect, and inhumane treatment inside the U.S. meat industry*, Prometheus.
- Ferriss, A.: 2002, Religion and the quality of life, *Journal of Happiness Studies* **3**, 199–215.
- Fitzgerald, A.: 2006, *Spill-over from "The Jungle" into the Larger Community: Slaughterhouses and Increased Crime Rates*, PhD thesis, Michigan State University.
- Foot, P.: 1967, *Theories of Ethics*, Oxford, chapter Moral Beliefs.
- Foot, P.: 2003, *Natural Goodness*, Oxford University Press, USA.
- Frank, D.: 1990, Anger as a Vice: A Maimonidean Critique of Aristotle's Ethics, *History of Philosophy Quarterly* **7**(3), 269–281.
- Frontline: 1999, The long walk of nelson mandela: The prisoner, <http://www.pbs.org/wgbh/pages/frontline/shows/mandela/prison/bam.html>.
- Frye, M.: 1983, *The Politics of Reality*, Crossing Press.
- Gauthier, D.: 1992, Artificial virtues and the sensible knave, *Hume Studies* **18**(2), 401–427.
- Gibbard, A.: 1999, Morality as consistency in living: Korsgaard's kantian lectures, *Ethics* **110**, 140–164.
- Glassen, P.: 1957, A fallacy in aristotle's argument about the good, *Philosophical Quarterly* **7**, 319–322.
- Harman, G.: 2000, The nonexistence of character traits, *Proceedings of the Aristotelian Society* **100**(2), 223–226.

- Herman, B.: 1993, *The Practice of Moral Judgement*, Harvard University Press.
- Hume, D.: 1998, *An Enquiry concerning the Principles of Morals*, Oxford University Press, USA.
- Hume, D.: 2000, *A Treatise of Human Nature*, Oxford University Press, Oxford ; New York.
- Hurka, T.: 2001, *Virtue, Vice, and Value*, Oxford University Press.
- Hursthouse, R.: 2002, *On Virtue Ethics*, Oxford University Press, USA.
- Isen, A. and Levin, P.: 1972, Effect of feeling good on helping: Cookies and kindness, *Journal of Personality and Social Psychology* **21**, 384–388.
- Johnson, H. M.: 2003, Unspeakable conversations, *New York Times Magazine* **February 16th**, 50ff.
- Kamtekar, R.: 2004, Situationism and virtue ethics on the content of our character, *Ethics* **114**, 458–491.
- Kant, G. t.: 1996, *The Metaphysics of Morals*, Cambridge University Press.
- Kant, P. t.: 1964, *Groundwork of the Metaphysic of Morals*, Harper Torchbooks.
- Keller, S.: 2007, Virtue ethics is self-effacing, *Australasian Journal of Philosophy* **85**(2), 221–231.
- Kesebir, P. and Diener, E.: 2008, In pursuit of happiness: Empirical answers to philosophical questions, *Perspectives on Psychological Science* **3**(2), 117–125.
- Korsgaard, C.: 2008, *The Constitution of Agency*, Oxford.
- Lebrecht, N.: 1985, *The Book of Musical Anecdotes*, Free Press.

- Lorde, A.: 1984, *Sister Outsider*, Crossing Press, chapter The uses of anger: Women responding to racism.
- Mackie, J. L.: 1980, *Hume's Moral Theory*, Routledge, London ; New York. J.L. Mackie.; Includes bibliographical references and index.
- Mencius, L. t.: 2004, *Mencius*, Penguin Classics.
- Milgram, S.: 1974, *Obedience to Authority*, Harper & Row.
- Monk, R.: 1990, *Ludwig Wittgenstein: The Duty of Genius*, Macmillan.
- Montaigne, S. t.: 2004, *On Friendship*, Penguin Books - Great Ideas.
- Nagel, T.: 1986, *The View From Nowhere*, Oxford University Press.
- Nakamura, J. and Csikszentmihalyi, M.: 2002, *The Handbook of Positive Psychology*, Oxford University Press, chapter The Concept of Flow, pp. 89 – 105.
- Narayan, U.: 1988, Working together across difference: Some considerations on emotions and political practice, *Hypatia* **3**(2), 31–47.
- Niven, L.: 1970, *Ringworld*, Ballantine Books.
- Niven, L.: 1992, *Playgrounds of the Mind*, Tor, chapter The Soft Weapon, pp. 180–240.
- Nussbaum, M.: 2006, *Frontiers of Justice*, Belknap.
- Oakley, J.: 1996, Varieties of virtue ethics, *Ratio: An International Journal of Analytic Philosophy* **9**(2), 128–152.

- Poloma, M. and Pendleton, B.: 1990, Religious domains and general well-being, *Social Indicators Research* **22**, 255–276.
- Postema, G. J.: 1988, Hume's reply to the sensible knave, *History of Philosophy Quarterly* **5**, 23–40.
- Prichard, H.: 1995, *Twentieth Century Ethical Theory*, Prentice Hall, chapter Does Moral Philosophy Rest on a Mistake?, pp. 37–47.
- Railton, P.: 1984, Alienation, consequentialism, and the demands of morality, *Philosophy and Public Affairs* **13**(2), 134–171.
- Rand, A.: 1964, *The Virtue of Selfishness: A New Concept of Egoism*, New American Library.
- Rawls, J.: 1971, *A Theory of Justice*, Cambridge, Mass.: Harvard University Press.
- Sabini, J. and Silver, M.: 2005, Lack of character? situationism critiqued, *Ethics* **115**(3), 535–562.
- Schueler, G. F.: 1999, Why is modesty a virtue?, *Ethics* **109**, 835–841.
- Shakespeare: 2001, *Othello*, Penguin Books.
- Sinclair, U.: 1906, *The Jungle*, Penguin Classics Deluxe Editions.
- Slote, M.: 1995, Agent-based virtue ethics, *Midwest Studies in Philosophy* **20**, 83–101.
- Solomon, D.: 1988, Internal objections to virtue ethics, in P. French, T. Uehling and H. Wettstein (eds), *Ethical Theory: Character and Virtue, Midwest Studies in Philosophy XIII*, University of Notre Dame Press, pp. 428–441.

- Spelman, E.: 1989, *Women, Knowledge, and Reality*, Unwin Hyman, chapter Anger and Insubordination.
- Stocker, M.: 1997, The schizophrenia of modern ethical theories, in R. Crisp and M. Slote (eds), *Virtue Ethics*, Oxford University Press.
- Street, S.: 2009, In defense of future tuesday indifference: Ideally coherent eccentrics and the contingency of what matters, *Philosophical Issues* **19**, 273–298.
- Sumner, L.: 1998, Is virtue its own reward?, *Social Philosophy and Policy* **15**(1), 18–36.
- Swanton, C.: 2003, *Virtue Ethics: A Pluralistic View*, Oxford University Press, USA.
- Taylor, G.: 2006, *Deadly Vices*, Oxford.
- Tessman, L.: 2005, *Burdened Virtues: Virtue Ethics for Liberatory Struggles*, Studies in Feminist Philosophy, Oxford University Press, USA.
- Tiberius, V.: 2006, Well-being: Psychological research for philosophers, *Philosophy Compass* **1**(5), 493–505.
- Toner, C.: 2006, The self-centeredness objection to virtue ethics, *Philosophy* **81**, 595–617.
- Unger, P.: 1996, *Living High and Letting Die*, Oxford.
- Wendell, S.: 1996, *The Rejected Body*, Routledge.
- Williams, B.: 1973, *Problems of the Self*, Cambridge University Press, chapter The Idea of Equality, pp. 230–249.

- Williams, B.: 1981, Persons, character and morality, *Moral Luck*, Cambridge University Press.
- Williams, B. and Smart, J.: 1973, *Utilitarianism: For and Against*, Cambridge, chapter Integrity, pp. 108–117.
- Wolf, S.: 1982, Moral saints, *Journal of Philosophy* **79**, 419–439.
- Wolf, S.: 1997, Happiness and meaning: Two aspects of the good life, *Social Philosophy and Policy* **14**, 207–225.