

LOVING THE BAD & NOT GIVING A DAMN:
A DEFENSE OF PSYCHIC DISHARMONY

Vida Yao

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy.

Chapel Hill
2016

Approved by:

Susan Wolf

Ryan Preston-Roedder

Geoffrey Sayre-McCord

Thomas E. Hill Jr.

Douglas MacLean

© 2016
Vida Yao
ALL RIGHTS RESERVED

ABSTRACT

Vida Yao: Loving the Bad & Not Giving a Damn: A Defense of Psychic Disharmony
(Under the direction of Susan Wolf)

We tend to think that it is better that one loves what one believes good, and hates what one believes bad. Similarly, we tend to think that it is a mark of virtue or rationality that one does what she judges she ought to do, and refrains from doing what she judges she ought not do. I argue that these “akratic disharmonies”, which occur between one’s evaluative judgments and one’s emotional or motivational responses, need not indicate her irrationality or a defect in her character, and that in fact, that her mind is disharmonious may be a sign that she is fully sensitive to the features of her situation. Moreover, her willingness or ability to sustain certain kinds of akratic disharmony can be the basis of underappreciated virtues such as graciousness, humility, self-esteem, and solidarity with her fellow human beings. Not only, then, is it *possible* for one’s mind to be akratic in the ways that I have described, there may be goodness in knowingly desiring, loving, or intending the bad.

For Chris

ACKNOWLEDGMENTS

A philosopher who visited UNC for one of our annual meta-ethics conferences, in sympathy with some of the unhappiness I was experiencing at the time, told me that graduate school had made her into a narrow sort of person, with rigid interests and concerns. However, once she finished up she found that her old self returned, her old passions now complemented by her philosophical interests rather than drowned out by them. I doubt that it occurred to her that this was exactly what I needed to hear at the time: I had the growing sense of having lost touch with other things, and in particular, other people, whom I loved. The thought that this was a sign of things to come was intolerable. But I held out, and she was right. So first and foremost, I want to express my deep gratitude toward my friends and family members. Each has served as sources of support, inspiration, and joy, even from a distance.

At UNC, I would not have finished without the friendship of Ben Bagley, Rob Smithson, Kate Nolfi, and Mary Renee Lindsay (who gave me her copy of *Gilead*). Special thanks are owed to John Lawless and Sam Reis-Dennis for their extensive discussion and comments on earlier versions of each of these chapters, and – more importantly – for their constant philosophical (and un-philosophical) company. My biggest debts are to my committee, and especially to my advisor, Susan Wolf. Not only did each member of my committee encourage and enhance my philosophical development, they also had faith in a project that was of great personal importance. They are not just teachers and mentors, but personal and philosophical inspirations deserving of both admiration, and affection. I cannot believe my good luck.

TABLE OF CONTENTS

1: Introduction.....	1
2: Grace & Graciousness: Goodness in Loving the Bad.....	6
I. Grace	9
II. Graciousness	23
3: The Undesirable & the Adesirable.....	38
I. What is the Guise of the Good Thesis?	41
II. An Illuminating Objection	50
III. The Human Being & The Human Good	53
IV. The Desirable, the Undesirable & the Adesirable	58
4: Boredom & the Disharmonious Mind	72
I. Boredom & Psychic Disharmony	72
II. Wholehearted Activity	75
III. Virtues & Vices of the Easily Bored	86
IV. Akratic Grass-counting.....	92
5: Fearsome Goods	101
I. The Claims	103
II. The Other Direction.....	117
6: Conclusion	127

I. An Adequately Complex Psychology	129
II. Plural & Conflicting Virtues	134
III. Virtues & the Importance of Evaluative Knowledge	137
REFERENCES	141

Introduction

If I abstain from fun and such,
 I'll probably amount to much,
 But I shall stay the way I am,
 Because I do not give a damn.

— Dorothy Parker, “Observation”

I will here focus on a familiar class of ways in which one’s soul can be disharmonious. These are cases in which one fails to feel, desire, or act, in accordance with one’s judgments or beliefs of value. Throughout, I’ll refer to this sort of disharmony as “akratic” disharmony. And I will be responding to two longstanding philosophical claims about akratic disharmony: the first we could characterize as *ethical* or *normative*, and the second we could characterize as *descriptive*.¹ Both claims reflect the basic philosophical tendency of venerating internal harmony over internal discord.

According to the ethical or normative characterization of the claim, it is a paradigmatic sign of irrationality, or of a bad human life or character, that one experiences this sort of akrasia. For Aristotle, the virtuous agent is not only able to get herself to perform the action she knows would be best to perform, she does so *wholeheartedly*, experiencing no conflict at all between

¹ This is likely to be too neat of a way of thinking about these claims – and indeed, I will here argue against both in a way that in some respects elides this otherwise helpful way of distinguishing the claims in question. But it is useful, at the outset, to begin with this way of organizing the kinds of claims that have been made about *akrasia*.

her judgments of what would be best, and her motivations to do it. For a contemporary expression of this ethical perspective, take for example Michael Stocker's observation that:

One mark of a good life is harmony between one's motives and one's reasons, values, and justifications. Not to be moved by what one values – what one believes good, nice, right, beautiful, and so on – bespeaks a malady of the spirit. Not to [believe good, nice, right beautiful and so on] what moves one also bespeaks a malady of the spirit.²

The idea of a unified or harmonious soul as an ethical ideal has a number of motivations speaking in its favour: the psychic painfulness of struggling to respond and act in a unified or simple manner; the lack of reliability in the performance of the best action when one struggles to do it; or the thought that there is something about integrity, or sincerity, that demands a person to feel in ways consistent with how they evaluate. Indeed, many of the struggles that any person may face in her life simply have to do with living up to the evaluative judgments that she makes, and we level charges of hypocrisy or duplicity should we find that a person claims to value something but fails to *do* anything about it.

Distinct from this ethical claim is another. This position can also be traced back to the ancients, but can – at least without elaboration – be harder to see the appeal of. Indeed, it has a tendency of either seeming clearly false or trivially true. According to this claim, unification is not a rational or ethical standard, but is a description of human psychology or action. In several dialogues, Socrates seems to claim that it isn't actually *possible*, in a clear-eyed manner, to fail to be moved by what one believes good, nice, right or beautiful and so on.³ In other words, clear-eyed akratic action is impossible: when it appears as though one is doing or desiring something

² "The Schizophrenia of Modern Ethical Theories", 453. To be fair, Stocker says that it is just "one mark" of a good life that one is harmonious in this way. In his later work, the claim is stronger: it is the sign of a non-defective character that one's soul is harmonious.

³ And as will be mentioned in the dissertation, claims of this sort are now taken as axiomatic in much meta-ethical discourse, especially by expressivists or quasi-realists. "Realists", on the other hand, maintain the connection between one's evaluative beliefs and one's desires or emotions by claiming that it is a standard of rationality that they be unified, and so might be better understood as offering a normative claim about psychic harmony.

that one knows is bad or bad to do, one in fact either lacks the relevant evaluative knowledge, or she has lost control of herself and so fails, in some sense, to be an agent.

One reason why this position is theoretically attractive is that it seems commonsense or truistic to think that our desires or our positive motivations and emotions generally, are responses to what we simultaneously take to be good in some way. To desire something *just is* to think it good, and so one cannot coherently desire something that isn't good at all, and would presumably desire the most what they think best.⁴ Such attitudes are what I will call "value-centric": should they be intelligent responses to the world at all (and so able to give rise to full-fledged action and other forms of rational response), they must be responsive to goodness.

There are ways to allow for certain kinds of disharmony *without* giving up the value-centric understanding of our positive attitudes. For example, if values are irreducibly plural, then somebody who has beliefs or motivations that reflect this plurality will be, in some sense, disharmonious.⁵ She may not know whether or not it would be all things considered best for her to volunteer at a soup kitchen, or work on her novel, or hang out with a friend; or she may know that it would be best to do one, but still feel emotional resistance to having to give up on the

⁴ Take, for example, how Donald Davidson summarizes Stuart Hampshire's denial of the possibility of akrasia in "How is Weakness of Will Possible?"

P1. If an agent wants to do x more than he wants to do y and he believes himself free to do either x or y, then he will intentionally do x if he does either x or y intentionally.

P2. If an agent judges that it would be better to do x than to do y, then he wants to do x more than he wants to do y.

P1 and P2 together obviously entail that if an agent judges that it would be better for him to do x than to do y, and he believes himself to be free to do either x or y, then he will intentionally do x if he does either x or y intentionally. (p. 28).

Davidson takes P1 and P2 to be "self-evident", and so rejects the thought that the puzzle about akrasia is a pseudo-problem.

⁵ This is how David Wiggins, for example, argues for the possibility of akrasia in, "Weakness of Will, Commensurability, and the Objects of Deliberation and Desire".

other. The disharmony in her soul, we might think, simply mirrors the disharmony of value in the world.

But the difficulty and tensions one might feel in negotiating plural and conflicting values is not the sort of disharmony that I will be interested in here. Rather, the sort of akrasia that I will be interested in here involve cases in which a person loves or desires or does something in ways that come apart from her beliefs about what things are good or worth doing. I will provide here a way of understanding how such cases are coherent and psychologically possible, by encouraging us to see the ways in which our motivations and emotions can be justifiably felt because responsive to the evaluative facts of one's situation. Importantly, however, those facts themselves need not be identical to considerations of goodness or badness of any kind. It has, in other words, been a mistake to think that the domain of the evaluative is limited to qualities of goodness and badness even pluralistically understood. We should reject, or place substantive limits on, the goodness-centric assumption that has guided our thinking about our positive attitudes.

I will also respond to the evaluative, or normative demand for unification. I will argue that there need not be anything significantly bad about a person should she be a clear-eyed akratic. And importantly, there need not be anything wrong with her even when her evaluative beliefs are *true*. So I will also not be discussing cases in which a person has mistaken beliefs about her situation. In those cases, we may just think that it's a good sign of her character that she simply doesn't feel the pull of the mistaken values that she holds. So for example, while it is true that Huckleberry Finn had false beliefs about what it is that he morally ought to do, we think the akrasia that he experiences when he finds he cannot turn Jim in to Miss Watson is a sign that

he has a good character, after all.⁶ But while sometimes we turn out to be *right* in acting against our best judgment, other times we still end up doing something downright stupid. Even in these sorts of cases, I will argue, there may be goodness in knowingly loving, desiring, or intending the bad.

⁶ As Nomy Arpaly has argued in “On Acting Rationally Against One’s Best Judgment”, if Huck does turn Jim in – even though it is against his best judgment – we might nonetheless evaluate him as doing the right thing, and so deserving of praise.

Grace and Graciousness: Goodness in Loving the Bad

I. INTRODUCTION

In the “Idea of Perfection”, Iris Murdoch provides a compelling image of a private moral improvement. A mother, M, feels hostile toward her daughter-in-law, D. M believes that that D, though good-natured, is “insufficiently ceremonious, brusque, sometimes positively rude, always tiresomely juvenile”.⁷ However, as time passes, M comes to question how she sees D and decides to look again. After reflecting and observing not only D but herself as well, M’s view gradually changes: D is discovered to not be “vulgar but refreshingly simply, not undignified but spontaneous, not noisy but gay, not tiresomely juvenile but delightfully youthful, and so on”.⁸ M’s hostility, we imagine, blossoms into something else. And her progress is a powerful ideal. We tend toward impatience, narrow-mindedness, and self-absorption; we have difficulties recognizing let alone overcoming aspects about ourselves that limit our ability to see others, as Murdoch puts it, “justly”.

Nonetheless, one might wonder what would have happened had M looked again, but found that her view did not alter – and not because of any failing on her part. What if, in short, D truly *is* brusque and positively rude? What then? Would M return to her “hardened sense of grievance,” now simply feeling more justified, and more hardened, than before? Murdoch herself

⁷ Murdoch, “The Idea of Perfection” in *The Sovereignty of Good*, p. 17.

⁸ *ibid*, pp. 16-17.

does not entertain this possibility. And indeed, some might insist that to ask these sorts of questions is to already have a misguided understanding of the ideal that she is articulating. The point of seeing another in this loving, admirable way is not to be *accurate*, but to be moved to find the best in another, or recast her badness in a more positive light by adopting a more pluralistic understanding of the ways in which a person can be good: to be *just*, in a very particular sense of the word. Those who worry that this could lead someone into, as Murdoch suggests, “deluding herself”, are perhaps themselves already too cynical or severe to appreciate this ideal in the first place.

But though this response is attractive, others may find it incomplete. In regards to cases that may be rare but not impossible, the questions I will take into consideration here could be put in the following manner: Is it possible for a person with a truly bad character to be loved, when careful attention would simply bring into sharper focus the ways in her character is flawed, even in light of a wide-ranging and flexible plurality of value? And, even if we can imagine such an attitude, would it simply be the sort of tragic, pathological affection that can attach hostages to their captors, or lovers to their abusers? Would it simply be perverse? These questions need not arise only because one is unwilling to see others in a rosier light. Rather, for some it arises when considering their *own* characters, and the limits of the love of others.

I shall here argue that the answer to the first question is yes, and to the latter two, no. But to reach these conclusions, we must reject a common way of thinking about our emotions and motivations, in general, and love in particular. There is a tendency, tracing back to Socrates and Plato but widely assumed in both historical and contemporary philosophical literature, to think of these attitudes in a way that is *goodness-centric*: that is, to think that if such responses are

intelligent responses at all, they are responses to, or projections of, goodness.⁹ Sometimes this thought takes the form of explicit moral, or rational standards: that such attitudes are rational or good to the extent that they are directed toward the good; irrational or bad to the extent that they are directed toward the bad.¹⁰ Should a positive emotional or motivational response be insensitive to goodness, it must not really be sensitive to anything at all: it would be “blind” or arational, like a craving or compulsion.¹¹

But there is a particular attitude that, though puzzling, should not be understood in either of these ways. That attitude is *grace*: a form of love that is not exclusively responsive to the goodness of its objects and may even be enhanced by, or even grounded in, the qualities of the beloved that the lover simultaneously recognizes are *bad*. As I will argue, this is because grace – of the form I will discuss here – is sensitive not to considerations of goodness, but to manifestations of *human nature*, including aspects of that nature that are bad. This feature of grace, as well as its distinct phenomenology, renders it importantly different from other attitudes that seem similar, such as forgiveness or faith, as well as other forms of love that can be sustained in awareness of the bad, but are ultimately projections or appraisals of goodness.

⁹ The first view is sometimes expressed as the thought that all desires and intentional actions are *sub specie boni*, or under the guise of the good; I will address this position directly in Chapter 3. It is also what is assumed by the thought that the *desirable* or the *lovable* is synonymous with the good. A more general psychological view is Socratic intellectualism, which treats our positive emotional and motivational states as responses to goodness, thus leaving no room for the possibility of *akrasia*. In contemporary work, David Velleman argues that love is a response to the *value* of the beloved in “Love as a Moral Emotion”. In *Taking Ourselves Seriously and Getting it Right*, Harry Frankfurt argues for a “projectionist” understanding of love, which takes love to be fundamentally a matter of *bestowing* value to the beloved through one’s love rather than appraising it.

¹⁰ Gwen Bradford has recently christened the principle of this position *amare bonum bonus*: to love the good is good (to hate the bad is good, love the bad is bad, and that to hate the good is bad). Thomas Hurka is perhaps the most prominent contemporary defender of this second claim. Both Hurka and Bradford suggest that it can be found in the work of Aristotle, Brentano, Moore, and Nozick. And as mentioned in the Introduction, Michael Stocker, who notably defends the *intelligibility* of desiring the bad, elsewhere argues that it is a mark of a good life that one’s motivations and evaluations are in harmony with one another, and a sign of a “spiritual malady” should they not be so.

¹¹ For example see Gary Watson, “Free Agency”.

After characterizing grace and providing a conception of it that renders it both coherent and I hope familiar, I will then argue that one's ability or willingness to sustain grace toward another person can be the mark of his possession of a set of virtues, including *graciousness* – a virtue in its own right – as well as humility, a deep-seated and informed self-esteem, resilience, and solidarity with one's fellow human beings. Correspondingly, a person's inability or unwillingness to experience or maintain such an attitude can mark his possession of certain vices, such as arrogance, a lack of self-esteem, an excessive demandingness, and a hostile or uneasy rigidity or harshness that may even take the form of misanthropy. The possibility of grace, and the virtues that can underlie one's experience of it, stand in tension with a tradition that has long emphasized the good of only and always loving the good. I will argue that not only is there sense to be made of such an endeavor; there is goodness in knowingly loving the bad, as well.

II. GRACE

I will use the term *grace* in a different sense than may be usual. Most notably, I will characterize an attitude that can be understood in abstraction from any theological commitments – one that can be recognized and experienced in strictly human circumstances.¹² Moreover, the kind of attitude I will discuss here is not identical to a form of *universal love* for all human beings that is sometimes upheld as a moral ideal. And though there is reason for treating grace as an aspect of all forms of love, and not a distinctive form a love in its own right, I will use the term in this latter way. This last point will mark my usage as distinct from, for example, Robert Adams' use of the term. He writes that grace is:

¹² In *Forgiveness and Love*, Glen Pettigrove also provides a conception of what he calls a “moral”, rather than theological conception of grace. I will note below how the understanding of grace I am offering here differs from his.

...love that is not completely explained by the excellence of its object. Within certain parameters, at any rate, it is not proportioned to the excellence of its object, nor conditioned on the degree of that excellence. To the extent that that degree can be measured, grace typically outruns it.¹³

As Adams observes, when it comes to the love of friends and family, for example, love that was completely explained by, or strove to be perfectly proportionate to, the excellence of the beloved would be perverse, if even a form of love at all. In light of this consideration, he suggests that grace is a quality of all *genuine* forms of love. While I have no objection to this use, I will here use the term to pick out a particular subset of forms of love that, though genuine, are less familiar and more peculiar. Attitudes of grace, as I will discuss them, are peculiar precisely because they are not simply *disproportionate* with the excellence of their objects, but because they are not always grounded, explained, or conditioned on excellence or goodness at all, but something else instead.¹⁴

This idea, especially given a philosophical tradition that emphasizes the good of loving the good, is no doubt apt to sound perplexing, if not straightforwardly incoherent. Take, for example, a case that is the inverse of M and D, in which one has affection for somebody, E, who appears to have good qualities of character. Imagine that upon carefully attending to E, you come to realize that what you took to be a sign of his benevolence was actually a sign of a calculating selfishness, and that what you took to be a sign of his love for his children was actually a sign of his need for control. It seems that you've lost exactly what grounded your affection for E – you may even wish that you had never learned the truth. With this sort of case in mind, how can we

¹³ Robert Adams, *Finite and Infinite Goods: A Framework for Ethics*, p. 151.

¹⁴ In “Must God Create the Best?” Adams suggests two different ways of understanding grace. First, that grace is a form of love that is not dependent on the merit of the person loved. Second, that the gracious person loves without worrying whether the person he loves is worthy of his love, or that the gracious person sees what is valuable in the person he loves, and does not worry about whether it is more or less valuable than what could be found in someone else he might have loved. I will here explore the first possibility: that grace is not dependent on the merit (or goodness) of the person loved, but on something else.

understand the psychological and conceptual possibility of *maintaining* one's love, let alone having one's love *be enhanced* by the bad qualities that one is now painfully aware of?

It would be best for us to begin with a less typical example, and consider the resources that a goodness-centric tradition has to illuminate it, before proposing how a particular understanding of grace can. Consider a scene from Marilyn Robinson's novel *Home* in which the main character, Glory, reflects on the quality of her brother's soul while helping him wash up from a night of heavy, self-destructive drinking.

Glory said, "You might rub your hands with shortening. That would probably dissolve the grease..." She took the can from the cupboard, scooped out a spoonful, and put it in his palm. She said, "Remember when you talked to me about your soul, about saving it?"

[Jack] shrugged. "I think you may be mistaking me for someone else."

"And I said I liked it the way it is."

"Now I know you're mistaking me for someone else." He did not look up from the massaging of his hands.

"I've thought about what I should have said to you then, and I haven't changed my mind at all... your soul seems fine to me. I don't know what that means either. Anyway, it's true."

He said, "Thanks, chum. But you don't know me. Well, you know I'm a drunk."

"And a thief."

He laughed. "Yes, a drunk and a thief. I'm also a terrible coward. Which is one of the reasons I lie so much."

She nodded. "I've noticed that."

"No kidding. What else have you noticed?"

"I'm not going to mention vulnerable women."

"Thanks," he said. "Very generous in the circumstances."

She nodded. "I think so."

He said, "I am unaccountably vain, despite all, and I have a streak of malice that does not limit itself to futile efforts at self-defense."

"I've noticed that, too."

He nodded. "I guess there's nothing subtle about it."¹⁵

This scene marks an important moment in the novel: an affirmation of how Glory feels about her brother, Jack, who she has slowly grown to know after a long estrangement. Earlier in their re-acquaintance Glory had remarked that she, as she puts it, *likes his soul the way it is*. Given the

¹⁵ *Home*, pp. 287-288.

sort of person Jack is, and the person *Glory* is¹⁶, she was puzzled by this; and at this point in the novel, her perplexity has not subsided. Instead, it is now accompanied by a certainty that simply compounds it. Indeed, she is expressing her sentiment explicitly within a context in which Jack has, once again, failed to live up to standards they both endorse. Though he tries to deflect the comment by suggesting that Glory doesn't really know what his soul is like, the rest of the passage reveals that her affection toward him is paired with a discerning awareness of his vices and the distinctive ways in which they manifest in his behaviour.¹⁷

Glory is similar to M in certain respects. She is self-critical, well intentioned, and able to see beyond her own prejudices when she attends to Jack. But their assessments of the objects of their attention travel in different trajectories. Initially, Glory *does* see Jack in an optimistic, charitable light: for example, as pensive rather than cowardly, or as playful rather than sharp-tongued. But as she attends more to Jack (with *his* encouragement to see him as he really is¹⁸) his flaws are simply brought into sharper relief. Nonetheless, her affection for Jack remains stable, and if anything, has been enhanced.

In these respects, her reaction is also notably different from perspectives that are offered within the novel of other ways in which attentive and even well meaning people might react to somebody like Jack. These responses are less puzzling than Glory's, as they are more in line with the goodness-centric understanding of love I described above. For example, Jack sees his own

¹⁶ Glory is both notably morally good, and pious; we have also learned by now that one of Jack's greatest wrongs – fathering and abandoning a child – deeply affected Glory, who grew to love the child before she died of an infection. We also learn that Glory's dreams of a simple family life had been dashed by a man who, like Jack, took emotional and monetary advantage of a “vulnerable woman”.

¹⁷ The scene continues, with Jack testing the limits of her affection by elaborating on what he calls his “mortal sins”. He clarifies to Glory that for some time of their estrangement he was in prison, not jail, as he had her previously believe. The scene concludes with her response: “I don't care if you were in prison.”

¹⁸ When he asks her to reconsider how she feels about the sort of person she is, to really *look at him* and what he has done, she responds: “What can I say? You're my brother”; to which Jack responds, dissatisfied, “You're not being kind”. This has bearings on an interpretation I consider below: about whether her love is just the kind of love that one has for people because of the particular relationship they bear to one another.

flaws with a clear-eyed shrewdness, and as such a view would predict, has no love to give. He suffers from deep bouts of regret and self-loathing; he is destructive and self-destructive in response. And Glory's reaction to Jack also differs importantly from their father's reaction to his Prodigal son. His love is one that veers between affection accompanied by excuses and minimizations of Jack's flaws, and harsh condemnation when those flaws are brought back into mind.

Glory's attitude is also distinct from another that is sometimes discussed alongside grace – that of *forgiveness*. Forgiveness too, is a positive attitude that knowingly outruns the excellence of its object, given that it would be a person's *flaws* that he is being forgiven for.¹⁹ Indeed, some might think that forgiveness is necessarily an act of grace, or that grace is a particular species of forgiveness. But there are two important respects in which grace differs from forgiveness that will help us further characterize the former.

The first is the sort of status that forgiveness presumes, and grace – at least when felt by one human being toward another – does not. In forgiving another, one presumes that one has the authority to do so, where this may involve being the person who was wronged or harmed by him, who is the source of the other's obligations, or who simply stands in the right sort of relationship to him.²⁰ Unlike her father, as well as the other "pious folk" that Jack has come across, Glory

¹⁹ Forgiveness, of course, is typically understood as an attitude that we take toward actions performed by others, but I mean here to discuss a form of forgiveness that can focus on the quality of a person's character as well.

²⁰ This cursory discussion of forgiveness is meant only to serve as a heuristic. Nonetheless, contemporary discussions of forgiveness in contemporary Anglophone philosophy almost without exception treat forgiveness as presupposing that the wrongdoer do something (such as apologize, or compensate) in order to merit forgiveness. See Jeffrie Murphy, "Forgiveness and Resentment" and John Wilson, "Why Forgiveness Requires Repentance". In *Forgiveness and Love*, Pettigrove argues that such accounts fail to recognize that one can forgive for good moral reasons that have little to do with desert and or merit. I agree with Pettigrove that forgiveness is best when accompanied by grace, but that standard accounts are right to preserve the thought that forgiveness is connected to ideas of merit and desert: one might forgive another in an ungracious manner.

does not take herself to have this sort of role in Jack's life.²¹ However she conceives of their relationship, she is not predominantly or even minimally in the business of – as he puts it – “saving his soul”. Though she is already aware of the bad aspects of his character, she is taken aback when he first asks her if that is what she is trying to do: “Save your soul? Why would I do that?” And after Jack explains to her why she might, she replies, “I’m flummoxed. I’ll have to give this some thought”.²² She does not take Jack to unredeemable or hopeless; it is that it doesn’t occur to her to think about their relationship in this sort of way. Rather than looking “downward”, in the way that forgiveness does, her attitude is marked instead by a kind of egalitarianism.

This aspect of their relationship also illustrates the second way in which grace differs from paradigmatic cases of forgiveness: grace need not focus on the goodness or good qualities of its object at all – not even its *potential* good qualities. Like faith in another, forgiveness is typically *aspirational* in a way that can preserve the basic thought that it is good to have positive attitudes toward the good, including the potentially good.²³ Jack can be forgiven for his pettiness, for example, on the assumption that he will make sincere efforts at changing for the better, and the faith that he will eventually become a better person or see the light. But Glory’s love isn’t like this. Again, this is not because she thinks Jack is hopeless, but because she just does not

²¹ This is not to say that her disposition toward him is simply one of sanguine patience. She does get angry with him for particular things that he does – for example, she is initially furious at him for nearly killing himself in the episode she is helping him clean up from. But for other offenses that do not involve her, but which she does recognize as genuine moral failings, she does not take herself to have authority to *forgive* Jack – some of these are things that only the particular person he wronged (for example, his wife) would have authority to forgive him for.

²² *Home*, 104.

²³ See Ryan Preston-Roedder, “Faith in Humanity”. In *Home*, such faith is expressed by another character (Lila) as a way of thinking about Jack: she suggests that God’s grace is a matter of understanding that *everyone can change for the better*. Perhaps as a way of answering whether this is what explains Glory’s affection, Jack immediately go off the rails once again, drinking himself into oblivion.

conceive of their relationship as one in which she has the standing either to forgive him, or see him as an inchoate version of the better person he will eventually become.

Her love then, is not best understood on the model of forgiveness. But another plausible explanation of her love would focus on the value of her relationship with Jack, and not Jack himself.²⁴ After all, Jack is her *brother*, and this can be enough to make sense of one's love persisting in ways largely insensitive to the excellence of the beloved, given the value of one's shared history or relationship. But while this form of love is no doubt genuine and valuable, it does not seem to fully capture Glory's affection. Glory has other siblings, whom she all loves *as her siblings*. But upon reflecting on her affection for Jack, she notices certain striking differences that reaffirm the puzzle she faces. She contrasts her attitudes toward all three of her brothers to her love of her "irksomely mature and responsible" sisters, and also notes that, "fond as she was of [her two other brothers], it was hard to remember that they had ever seemed marvelous to her" in the way that Jack did, and does.²⁵

Perhaps it is in this last observation that we can find a goodness-centric explanation of Glory's love, after all. While Jack's character is composed out of various bad qualities that Glory can plainly see, when she assesses Jack *as a whole*, he achieves a kind of value that those bad qualities do not have on their own, nor the value they would have if simply added together. Jack's cowardliness, meanness, and vanity are bad qualities – but we should not assume that the goodness of *Jack* is something that can be determined by adding up the value of these parts. As

²⁴ Cf. Niko Kolodny, "Love as Valuing a Relationship".

²⁵ *Home*, 71. The particularity of her love for Jack also doesn't seem to be explicable by, for example, a universal love for humanity, and Jack as the person that she happens to direct her love toward in light of this more general love. But the question of what justifies love's particularity is a difficult one that I will not fully address here.

what G.E. Moore calls an “organic unity”, Jack is “marvelous” – and this is why Glory loves him.²⁶

Moore was no doubt right to emphasize the complexity of the relationship between the value of a whole and the value of its parts. And we do seem to think that some undoubtedly excellent people, are excellent in part because of some aspect of their characters that, on their own, are bad. Being short-tempered may be a bad trait, but when it partly forms the organic unity that is Nina Simone, we see it as a necessary part of an overall excellent being.

But while it is true that we do not, for example, determine the value of a piece of music or an excellent meal or a person by adding up the value of their constitutive parts, this answer only gets us so far. First, it leaves open the question of how it can be possible that some genuinely bad aspect of a person *could* enhance him in such a way that makes him overall, more lovable. Second, we might wonder what to do in cases in which, taken as a whole, the person *doesn't* add up to something that is actually marvelous. Perhaps Michael Jordan's excellence is partly constituted by his arrogance, or Emma Woodhouse's excellence partly constituted by her snobbery, or Margo Channing's excellence partly constituted by her petulance; but what about somebody like *Jack*? It is far from clear that *he* is actually good, let alone marvelous, even though he *seems* to be. This is precisely where Glory finds herself puzzled. And when we take the sort of question we began with seriously, and consider those people in our *own* lives who may stand in as *our* Jack, this answer depends on a possibility that simply seems incredible. So the question remains.

Maybe the answer lies in simply allowing for some forms of love to transcend justification or rationalization. These forms are arational, like emotional glosses that one

²⁶ Moore, *Principia Ethica*.

experiences, regardless of what one is presented with and regardless of whether one's response is warranted. After all, love, some think, is blind. And grace understood by Judeo-Christian traditions is precisely God's puzzling and perhaps arational undeserved and undeservable love. Perhaps we need to make peace with the idea that love in general, as Adams puts it, simply "outruns reason".²⁷ And maybe this is all there is to Glory's love, in particular.

But I believe that there is more to be said, at least about some cases of love. An answer is made available should we conceive of grace in a certain way: as both clear-eyed, and informed or enhanced by qualities that the lover simultaneously recognizes as bad. This answer helps illuminate both why a bad quality can make some excellent people seem even *more* excellent, and why we may be able to love those who are far from it.

This proposal will sound mysterious so long as we are still limited by the goodness-centric tradition, so accepting it involves putting certain constraints on that way of understanding the justificatory nature of our attitudes. My suspicion is that what this tradition tends to eclipse is that there are a number of different, heterogeneous qualities of objects that render our emotions appropriately or inappropriately felt, which can enhance or diminish those emotions, and which are not identical or reducible to considerations of *goodness* and *badness*. In advancing this point, I am setting certain limits on the "broad claim", voiced by Rosalind Hursthouse but implicitly accepted by others, that "the emotions involve ideas or images (or thoughts or perceptions) of good and evil, taking "good" and "evil" [or "value" and "disvalue"] in their most general, generic sense, as the formal objects of pursuit and avoidance."²⁸ Unless we are to read this claim as stipulative, rather than substantive, we should reject it.

²⁷ Adams, *Finite and Infinite Goods: A Framework for Ethics*, p. 173.

²⁸ *On Virtue Ethics*, p. 111. Hursthouse is responding to a point made by Bernard Williams. In "Morality and the Emotions", Williams suggests that the reason why relatively little attention is paid to the emotions in the moral

My ambition is not to propose an alternative, broad claim in response – this risks violating the spirit of a proposal meant to emphasize the heterogeneity of ways in which our attitudes may be justified. Rather, by focusing on particular emotions, we find that those attitudes are responsive not to *goodness* and *badness*, but to other evaluative qualities of their objects.

Here, I shall focus only on grace, and propose one possible explanation for why grace, in particular, can be appropriately felt without being grounded in or informed by the goodness of its object. Grace is sensitive to aspects of a person that, though they may be bad, are nonetheless distinctively *human*. Grace is perhaps best understood as a love for *human nature*, and expressions of that nature. One might initially characterize grace as a love for *humanity*, but importantly, it is distinct from the ideas attached to that expression such as the thought that human beings are infinitely or incomparably valuable, or the thought that particular actions must be taken in order to secure or respect their value. Moreover, it is not love for one particular feature of our nature such as our rationality, or our capacity to set ends, or some abstract metaphysical mark of humanity. Rather, it is a love for the distinctive, concrete or *phenomenal* ways in which human nature exhibits itself in a person's character and behaviour – including the bad ways.

This proposal, I hope, illuminates grace as something more familiar, and more familiarly experienced and expressed. Take for example, Angel Clare's first experience of Tess Durbeyfield's beauty:

philosophy of the time (1965), is because of an overly narrow concentration on highly general moral language such as the language of "good" and "right". He writes that such a focus "[helps] to push the emotions out of the picture. If you aim to state the most general characteristics and connexions of moral language, you will find not much to say about the emotions; because there are few, if any, *highly general* connexions between the emotions and moral language," (p. 208). He is not thereby endorsing, for example, non-cognitivism about the emotions – he clarifies that "there [is] no basic reason why a generous approach to the linguistic endeavor (i.e. richer, more specific evaluative language) should not have embraced those features of our speech about morality that reveal the parts played by the emotions," (p. 207). Hursthouse, in response, is suggesting that such highly general connections *do* exist; like Williams, I am skeptical.

He had never before seen a woman's lips and teeth which forced upon his mind with such persistent iteration the old Elizabethan simile of roses filled with snow. Perfect, he, as a lover, might have called them off-hand. But no — they were not perfect. And it was the touch of the imperfect upon the would-be perfect that gave the sweetness, *because it was that which gave the humanity*.²⁹

Or – for a more contemporary example – take comedian Ellie Kemper's answer when asked what her favourite joke is:

...Anytime anyone's umbrella turns inside out on a rainy, windy day — I can't stop laughing. It's like, who are we kidding? We're all just helpless babes in the face of nature. And we all just look so dumb.³⁰

While of course, neither a less than perfect smile nor a person's inability to keep her umbrella inside in on a rainy, windy day, are aspects of a human being that are *bad*, the general sentiment expressed by both of these thoughts can be extended to things that more plausibly are.

Return to Jack: he's a drunk, a thief, a coward, a liar, a rake; he's vain, and sometimes malicious, seemingly just for the sake of being so. While these are all flaws in a person's character, they are distinctively or characteristically *human* flaws. Other contenders for characteristically human flaws might include such things as laziness, vanity, brattiness, and particularly human ways of being stupid. Grace does not try to re-conceptualize these qualities as *good*, or as better than they are commonly thought to be – it rightly takes those that are bad *as* bad, but nonetheless also as the potential grounds for, or even enhancements of, its existence.³¹

²⁹ *Tess of the D'Urbervilles*, p. 119, my emphasis.

³⁰ Fitz-Gerald, "61 Comedians Recall Their Favorite, First, and Life-Changing Jokes".

³¹ In suggesting that aspects of human nature are what can "rationalize" grace, or render it appropriate, I do not mean to endorse the stronger thought that such aspects rationally *obligate* love. As Adams puts this point, "It is a prejudice to be avoided that reasons that tend to justify a belief or action or attitude, in the sense of vindicating its rationality, must also tend to compel it, in the sense of rendering it irrational to abstain from it," (*Finite and Infinite Goods*, p. 163). A person who did not experience the attitude of grace toward a manifestation of human nature is not being *irrational*, but he may be being ungracious.

This provides us with one way to understand Glory's particular affection for Jack. While their brother Teddy is, by all accounts, a much *better* person than Jack is, and though Glory loves both Teddy and Jack equally as her brothers, Teddy does not inspire the *same* sort of love that she has for Jack who – though worse – more fully exemplifies the full range of human nature.³² This proposal can also explain why people who, perhaps unlike Jack, are *excellent*, can be *more* lovable to us precisely because of their flaws. Why is it that Tess Durbeyfield's overall beauty is made "sweeter" by an imperfection, or that one's love for Nina Simone partly *depends* on her extreme and chaotic temper? On a goodness-centric approach to our attitudes, we may be tempted to revise our evaluative standards to make sense of these attitudes: such patterns in our affections suggest that these "flaws" are not really so, or that we do not really believe them to be so. Perhaps the lesson is that we need to recognize that there are legitimate standards and perspectives from which these aspects *are* genuine goods.

While it is true that we should be skeptical of our evaluative standards, and that we should also be flexible and pluralistic in our considerations of value, the understanding of grace I propose *supplements* such standards. It allows that some flaws – even when they are *genuine* flaws – can humanize their possessors in a way that can endear us to them: it reveals to us that they are not actually angels, or (for different reasons) machines, but just particularly excellent *human beings*.

Another way to further illuminate both the structure and the phenomenology of grace is to turn our attention to a set of attitudes that share the basic features that render grace puzzling, but with an added complication. These attitudes are particularly puzzling within a goodness-centric tradition because at a certain level of abstraction, such attitudes can be understood as a

³² In discussing their very good brother, Glory remarks, "Teddy's all right. He means well", to which Jack replies, "Despite his virtues and accomplishments". Glory considers this: "Yes, in a way that's true". (*Home*, p. 124).

love of badness for *badness's sake*.³³ I will here focus only on aesthetic examples of such attitudes, and common names for this sort of love is a love of *kitsch*, or *camp* – or more plainly, a knowing love of the tacky or the cheesy.

One response to such attitudes is to think that if they are coherent, it is because they are not actually attitudes of sincere affection or love. They are instead the result of a cynical and arrogant sort of meanness – a desire to seek out the bad that involves taking pleasure in the failures of others in a condescending, petty way. Maybe this is how we should understand those who enjoy watching films like *Showgirls* or *Battlefield Earth*, or who delight in sincere expressions of trite clichés: *I love you but I'm not in love with you; I thought I was there to teach the students, but soon realized that they were teaching me*. On this interpretation, people who enjoy such things are tickled by the thought that those involved in the production of these enterprises and expressions had no idea they were creating something so *awful*, indulging in the kind of smug satisfaction one might get from watching other people make fools of themselves. This sort of attitude ridicules not just the creator, but looks sideways as well: it indulges in the thought that other people are being duped.

This worry is illuminating, and it provides us with a legitimate explanation for why some people do desire to engage with badness for badness' sake. But it is not the *only* explanation: grace makes another available to us. On one standard understanding, kitsch and camp appeal because such objects or performances are a presentation of something that produces reactions in us that are *simulacra* of a real emotional engagement with something actually beautiful, or good.

³³ There has been relatively little philosophical discussion of such cases, perhaps because they are difficult to make intelligible within standard philosophical understandings of action and psychology. David Sussman is one exception: in "For Badness' Sake" he discusses these cases, arguing that such actions are in an important sense unintelligible, and that we have good reason to avoid their performance because of this. I do not think that they are unintelligible, and that even if they were, this alone would not be a reason to refrain from performing them. In *Value in Ethics and Economics*, Elizabeth Anderson also considers the idea that a person could knowingly enjoy kitschy art, but suggests that it is because attitudes of liking are *arational*.

The huge, pleading eyes of a Margaret Keane portrait, the schmaltziness of a pop song, or the mawkishness of a cliché tug at one’s emotions, not by way of presenting anything beautiful or good to the viewer, but by way of hijacking one’s weaknesses.³⁴

This illusory quality is one reason why some disparage kitsch – not simply for being bad, but for being *manipulative*. It tricks us in predictable ways by taking advantage of our shared aesthetic frailties. This is why it is often seen as “art” for the masses; for the lowest common denominator. However, in taking up a particular sort of stance toward kitsch and camp, one can *realize* that one is being deceived by such objects without thereby losing one’s affection for such things, or reacting with disdain for their creators, or condescension toward others. In doing so, one may be recognizing and acknowledging in herself a genuine affection for something, without at the same time endorsing or reinterpreting this affection as directed toward anything *good*. If she is mocking the object of her affection, its creator, or others who would enjoy it, she is mocking *herself* as well; it is delight paired with acknowledgment and love for the parts of our shared natures that are implicated in the deception. In other words, a knowing love of the bad in such cases need not always be the result of condescending elitist irony, but grace.³⁵

Finally, given this understanding of grace, we can see that it does have justificatory standards, and so, limits. It is unfitting or impossible to experience it toward beings that are distinctively inhuman, though they may be members of our biological species. Consider how Helen Prejean – the woman at the center of the book and film *Dead Man Walking* – characterizes

³⁴ As Roger Scruton has put this point, “Kitsch is fake art, expressing fake emotions, whose purpose is to deceive the consumer into thinking he feels something deep and serious, when in fact he feels nothing at all.” Scruton probably does not mean that the consumer feels *nothing*, but instead, that what he does feel is an illusion of something else. See Scruton, “A Point of View: The strangely enduring power of kitsch”.

³⁵ In “Notes on Camp”, Susan Sontag writes: “Camp is generous. It wants to enjoy. It only seems like malice, cynicism. (Or, if it is cynicism, it's not a ruthless but a sweet cynicism)... Camp taste is a kind of love, love for human nature,” (291-292). I hope that my discussion here further illuminates Sontag’s proposal. Thank you to Benjamin Bagley for encouraging me to clarify my thoughts about such cases.

her first visit to death row inmate, Patrick Sonnier. Prejean knows ahead of time that Sonnier has been convicted for the murder of a teenaged couple, and though she does not believe that he should be sentenced to death because she thinks it always unjust to do so, she does not doubt that he was indeed, responsible for this terrible crime. She writes of the encounter: “I am meeting him in a crucible, and I am surprised by how human, even likable, he is. Despite his friendly letters I had half expected Charles Manson – brutish, self-absorbed, paranoid, incapable of normal human encounter”.³⁶ While it is difficult to know how and where to draw the boundary between the human and the inhuman³⁷, we may nonetheless have a sense of what Prejean means when she draws a contrast between somebody who is recognizably human with a biological human being whose mind is so deranged or scrambled that we are not really able to understand him as having a human character with character traits, good *or* bad.³⁸

III. GRACIOUSNESS

Given this understanding of grace, we can now explore in more detail why it can be a *good* attitude for a person to exhibit, and why *graciousness* is a virtue. For some, it may be obvious why it is admirable to treat people better than they deserve, or to benefit others more than is merited. But why might this be so? Moreover, what is added to our understanding of grace if we see it as responsive to human nature, rather than simply that aspect of all forms of genuine love that outruns reason? These are two features of grace that can be discussed separately in light of understanding graciousness as a virtue: its content (that it is directed toward

³⁶ *Dead Man Walking*, p. 31.

³⁷ I will say more about this distinction in chapter 3.

³⁸ Aristotle, too, seems to make this sort of distinction when he discusses the category of the “beastly”, or the “brutish” in Book VII of the *Nicomachean Ethics*, suggesting that while brutish states are “less grave than vice”, they are “more frightening” (NE1150a37).

human nature) and its form (that it is a knowing love of the bad). Both aspects will help to illuminate why graciousness is desirable, as well as why, like other virtues, it is a corrective to certain natural tendencies.³⁹

Loving Human Nature

Grace has two features that render it among some of the most desirable forms of love. Consider again Glory's love for Jack. First, not only does Glory love Jack in the sense that she cares about his wellbeing, as well as in the way that siblings love one another as siblings, she also *likes* him. She is interested in his character in a holistic way that takes into account aspects of himself that are truly his, though truly bad. Second, her love is attentive and realistic about the kind of person Jack really is. It doesn't mask, reinvent, or try to justify or excuse.

While some might be satisfied with, or positively want forms of love that *lack* this second aspect, those of a more critical and self-reflective bent may find themselves in need of something else, as well.⁴⁰ Jack falls into this latter group. He feels uneasy around his father, even when his father is affectionate toward him, precisely because he knows that such warm feelings are based on an inaccurate or incomplete understanding of what he is like. When others' assessments of him are out of step with his own understanding of himself in this way, he experiences an isolating disorientation.⁴¹ In contrast, Glory can reflect and speak truthfully about Jack, but without the hostile attitudes that have made the truth so painful to him, and others. Because it both discerns his faults, and *loves* them as aspects of his character, Glory's love is one that is

³⁹ As suggested by Philippa Foot in, "Virtues and Vices".

⁴⁰ Of course, it also depends on who the lover is, and what sort of love is in question: one might positively want, for example, one's parents to always see the best in her, but not want this sort of love from her friends.

⁴¹ For example, Jack rejects Glory's attempts to excuse his past behaviour because of his youth: "No I wasn't young. I don't believe I ever was young. Excuses scare me, Glory. They make me feel like I'm losing hold" (p. 154).

directed toward Jack as he actually is in his entirety. And because of this, Jack can trust her love in a way that he finds unable to trust the love of others. In letting down his guard, he is given the rare chance to connect with another person who truly sees him as he is. This is one sense in which this kind of grace is analogous to its theological counterpart: it saves.

Given the content of grace, graciousness has important connections with other virtues, as well. Above, I suggested that grace is different from forgiveness because it presumes no *authority* over its object. Indeed, grace can be *inhibited* by a tendency to hold oneself in a higher regard than the beloved – to “love” the flaws of others, either as an expression of one’s will to power, or in the condescending way in which one might “love” something that really, one simply *pities* or finds ridiculous. To counteract this tendency, a fully gracious person is likely to be disposed to direct her attitude not only toward others, but toward herself as well. Her affection for human foibles and shortcomings can be paired with, and enhanced by, an awareness of her own comparable flaws: her delight, for example, with a person’s clumsiness is inextricably paired with a recognition of her own. She acknowledges that she is, after all, a human being as well.

This allows her to avoid the vices of condescension, arrogance, or self-righteousness. She instead manifests *solidarity* with other human beings in light of their shared nature – including the nastier and stupider bits of that nature. A gracious person does not think that the bad qualities of human beings are necessarily things to be *simply* rectified or eliminated; she does not suffer from the kind of hatred or loathing that some might experience when reflecting on or

encountering aspects of human nature that can be confining, or frustrating, or regrettable. She avoids, in short, the vice of misanthropy.⁴²

Moreover, because the fully gracious person is able and willing to recognize and love her own flaws as well as the flaws of others, she exhibits a form of self-esteem that doesn't depend on ignoring or being blind to her own faults. Having recognized their existence and nonetheless being able to love herself in a way that is informed by that recognition, she isn't easily embarrassed or ashamed when those flaws are made apparent or manifest in her behaviour, though not because she has found a way of convincing herself that they are really not so bad, after all.⁴³ Harry Frankfurt, in discussing the importance of "taking oneself seriously" by loving what one values wholeheartedly, he considers whether we can, after all, take ourselves *too* seriously. The advice he offers at the conclusion of his book: take yourself seriously, but if you cannot – "at least be sure to hang on to your sense of humor".⁴⁴ But a sense of humour – about oneself and about human nature, more generally – is no *consolation* prize, one to be relied upon when, desperately trying to close his umbrella on a rainy day, a person finds that he cannot live up to the demands of self-seriousness. It is good and worth aspiring to in its own right.

⁴² In a discussion on the defensibility of "speciesism", Bernard Williams mentions misanthropy as one explanation for the philosophical motivation of always holding human beings to the justificatory standards of some "higher court". He writes in "The Human Prejudice":

When the hope is to improve humanity to the point at which every aspect of its hold on the world can be justified before a higher court, the result is likely to be either self-deception, if you think you have succeeded, or self-hatred and self-contempt when you recognise that you will always fail. The self-hatred, in this case, is a hatred of humanity. (p. 152)

⁴³ There is a memorable scene from the recent film *Silver Linings Playbook* that illustrates the kind of confidence or self-esteem I have in mind, in which Tiffany rebukes Pat for a number of reasons (including calling her a "slut"), forcefully telling him, "There will always be a part of me that is dirty and sloppy, but I like that, just like all the other parts of myself". It is also the sort of self-esteem that Dorothy Parker exhibits in the poem with which I began this dissertation.

⁴⁴ In *The Reasons of Love*, Frankfurt denies the importance of knowledge in love: on his picture, love is just a wholehearted endorsement of whatever it is that one finds one *can* love. Such a picture cannot capture what I think is distinctive about Glory's love for Jack.

Knowingly Loving the Bad

Having explored one aspect of grace – that it is a love of human nature – let’s now turn to another – that it is a knowing love of the bad. Though knowingly loving the bad – any bad – is often taken to be a sign of a familiar sort of weakness, it can in fact take a concerted effort to experience, and maintain. There are three sorts of natural impulses that the gracious person would need to resist. First, there is a sense in which she would, after all, be justified in turning away from or hating aspects of herself or others that she thinks of as bad. Second, should she find that her affection of the beloved is stable she must also resist recasting the object of her attitude in a more positive light. And the third would be to strive to improve the object to better match one’s desires and positive emotions. It is these different impulses that account for the vacillation in Jack’s father’s attitudes: either he can only lament the flaws that he plainly sees in his son, or he loves in a way that is blind to them.

To illuminate the virtues connected with an ability to resist these impulses, it will be helpful to discuss more generally the sorts of virtues that are exhibited by those who knowingly love the bad, more generally. To guide this discussion, consider the following example:

David Chang is founder and former chef of (among others) *Momofuku Ko*, a restaurant that was awarded two *Michelin* stars in 2008. Like many very excellent chefs, Chang is often asked what his own diet is like. In a number of interviews in which this question comes up, his response is consistent: fried chicken fingers. Are these *good* fried chicken fingers? No, he ensures us, he is not so discriminating: “Any processed chicken from any place—I’ll order it in a heartbeat”.⁴⁵

⁴⁵ For an extended profile of Chang that illustrates (among other things) the aspects of his personality I am isolating here, see Larissa MacFarquhar’s “Chef on the Edge” in *The New Yorker*.

I begin with the simple observation that I, at least, find there to be something immediately charming or likeable about Chang, given this bit of information about his appetites, as well as certain background assumptions about his character. These assumptions include first, that he possesses knowledge about what kinds of food are really truly good and excellent along the many dimensions that are relevant to evaluating food. Second, that given his skill and his resources, he would be able to pursue those foods should he desire them. But, third, that in spite of this knowledge and capability he desires to eat food that he recognizes is *not* good: food that isn't good for him, is crude and bland, or garish, in flavor, and is (given exactly what goes into creating fried chicken fingers at most fast food places) even morally dubious. And he desires such things intrinsically: that is, for their own sake and not because they just happen to be cheap and easy to obtain.

Notice that we might react differently to somebody who didn't possess the relevant sort of evaluative knowledge: someone who desires fried chicken fingers because he hasn't had the opportunity to or the interest in discovering food of better quality and taste, who believes that fried chicken fingers from any fast food place simply *are* good, or who simply doesn't truck in the evaluation of food. We would also react differently to somebody who, though possessing the relevant knowledge but lacking intrinsic desires for chicken fingers, desires to eat them in certain circumstances only out of convenience or because of a lack of resources or ability to obtain something better instead. Perhaps in these sorts of cases, we might withhold judgment entirely, or our judgment will be sensitive to the considerations about how one's perspective arose.

And importantly, at least some of us would react differently to a person who possesses the relevant evaluative knowledge and resources, but whose desires are, unlike Chang's, fully "in line" with that knowledge: a person who has no intrinsic appetite at all for anything that isn't

good. In another interview, another excellent chef is asked whether or not she has ever tried an “Oreo, Twinkie, or gummy bear” just to see why it is that so many North Americans love these sorts of things, in spite of their obvious vices. She responds, “Oh, I have tried it. I mean I certainly tried it as a teenager. I have eaten at McDonald's once... I was surprised it didn't have taste from my point of view — and, well, I was in and out in five minutes”.⁴⁶ In contrast to this person who – at least in this interview – seems distant or detached, Chang seems decidedly likeable or relatable.

These observations alone are not all that illuminating when it comes to the question of what could be substantively good about a person who knowingly desires the bad, and what may be bad about a person who always prefers the good to the bad. But I think that these initial reactions are indicative of our recognition of deeper and more widely recognized virtues possessed by those who knowingly desire or prefer the bad to the good.

Take into consideration again how it is that a person might achieve more constancy between her evaluative beliefs and her desires. As suggested, there seem to be three ways for one to go about doing this. The first would be to change one's evaluative beliefs to better match one's desires. The second would be to try to change one's desires to better match one's beliefs. And the third would be to strive to change the object to better match one's desires and positive emotions.

The first approach risks taking too seriously that one's own desires or felt positive attitudes have a kind of authority over the value of something. To illustrate, consider an example provided by Simon Blackburn, in discussing the various ways in which we can interpret a person such that her evaluative beliefs are consistent with her desires. He asks us to imagine a person who no longer experiences pleasure in drinking wine, but who nonetheless declares that a particular bottle is “good”. Blackburn suggests that given that she has no interest in actually

⁴⁶ "Alice Waters on Talk to Al Jazeera." edited by David Shuster: Al Jazeera America, 2013.

drinking it, she cannot be sincerely judging that the wine is *actually* good. According to this person, Blackburn suggests, the wine “once merited” such a judgment, but no longer does, because it is no longer something that she enjoys.⁴⁷

But it seems that a sincere judgment of goodness should not be used simply as a way of reporting what one happens to enjoy, or desire. This isn’t a semantic point, nor is it one that engages directly with the expressivism that leads to Blackburn’s comments here. Rather, it’s an ethical one: to use “good” in the way that Blackburn suggests is intelligible, and indeed, in some contexts, we do sincerely mean to just report what it is that we’d desire or prefer. But depending on the context, it places too much epistemic authority on one’s own subjective preferences as determining, or as evidence of, the positive qualities of a thing. And when a person places too much emphasis on her own preferences or desires, it can be a sign of evaluative arrogance. The issue here isn’t simply a matter of unjustifiably assuming that one has a kind of authority that one doesn’t: even those who *are* experts about some domain, and who *are* justified in taking their judgments more seriously than the judgments of non-experts, would do well to consult something other than their own tastes when determining whether or not a given item within that domain is really good.⁴⁸

⁴⁷ *Ruling Passions: A Theory of Practical Reasoning*, (Oxford: Clarendon Press, 1998), 60. Unlike more extreme versions of non-cognitivism, Blackburn does not think that it is *impossible* for one’s beliefs about goodness and one’s “pro-attitudes” to diverge from one another, but he does think that these cases will necessarily be “out of joint” – that is, unusual enough to make us suspicious about their status. Indeed, it is important to his expressivism that this be the case. However, though such people might be *unusual*, this is not because they are impossible or in any way *ethically* suspicious.

⁴⁸ In “Goodness and Choice”, Philippa Foot argues against Hare’s prescriptivism by emphasizing how unusual it is to actually be in the sort of position where one’s choice *at all* determines the goodness of the thing in question. As she puts the point, to determine whether x is a good x is typically a matter of *discernment* and awareness of the standards relevant to assessing x’s. I am largely in agreement with her arguments there, and simply want to add that it may be a sign of arrogance should one fail to rely on such standards in one’s determination of the value of a thing. Foot, “Goodness and Choice,” in *Virtues and Vices: And Other Essays in Moral Philosophy*, (Oxford: Oxford University Press, 2002).

Consider then, the other way in which a person might attempt to eliminate his desires or affection for the known bad. Rather than re-assessing the quality of such things, he will try to modify his attitudes in light of his realization that they are directed toward things that are bad. However, depending on his motivations in doing this, he may exhibit another set of vices.

To illustrate, bring into mind what it is like to realize that something that you happen to sincerely enjoy *really* is pretty awful. Perhaps this is a particular song on the radio, or a television show, or (like with Chang) certain kinds of food. Now, one might try to *refine* one's tastes such that these things no longer appeal, and what *does* appeal now are only good or excellent music, reading material, and food. Or one would try to imagine up a certain perspective from which such a thing really *is* good, after all. But *why* would a person do either of these things? Why is she so concerned to only love the good, and to hate the bad? Is she unable to live with the idea that her motivations are not so pure? Is the possibility of this something that she is ashamed of, or embarrassed to discover in herself? Is she fixated on being immune from criticism? If *any* of these motivations ring true, then we should again remind ourselves of the comparative authenticity and self-esteem of her counterpart, who is able to simply enjoy what he enjoys, while simultaneously acknowledging that such things may very well be pretty lousy.

Now of course, sometimes it doesn't take effort to love the good rather than the bad. Sometimes loving the good and hating the bad can come quite easily – one begins to eat better food because she enjoys it, and eventually she just finds that she no longer enjoys what she once loved. So take into consideration a person whose positive motivational and emotional responses are wholly for goodness, not because she has actively *cultivated* these attitudes, but because it has just come naturally to her to love the good and hate the bad.

Imagine again Chang's counterpart, this time considering that her distaste for "it" (that is, Oreos, Twinkies, gummy bears and McDonald's) comes to her naturally. While we might appreciate and admire the knowledge and skill of this person – just as we admire Chang for his – some of us, at least, might notice that our responses toward the two differ beyond this admiration. Even though we might *agree* with this second chef that gummy bears and Twinkies truly are bad forms of food, and though we might be in awe not simply of her self-control (since she does not enjoy these things at *all*), but her *wholehearted* appetites for the good, the admiration or esteem that we might feel for her may be noticeably absent of the affection or warmth that we feel toward someone like Chang.

Consider too, that the more dimensions along which a person loves only the good and hates the bad, the more distant she may become, and the less affection she may inspire. Imagine a person who has excellent taste in music but who cannot bear sitting through drunken karaoke at the bar, or a person who has excellent taste in movies but cannot stand any that are stupid or banal. Even though we might agree with the evaluative judgments that this person makes, we might also – if we're at all inclined to be friends with her, which we might not be – want to encourage her to lighten up or take herself less seriously. Such a person is concerned with goodness in a way that can inhibit one's affection of or attraction to her. She may be rigid, inflexible, uptight, or stodgy (at least in American English).

Now, along with these two psychological tendencies, there is another aspect of our evaluative lives that can make it difficult to experience grace that is worth mentioning. Recognizing this aspect will reinforce this last point about friendship, as well as bring into view the third strategy one could adopt in order to achieve a more harmonious psychology: the strategy of, when possible, *improving* the objects of our affection so that they may become good

in a way that matches our love. And here, I want to discuss not just a less than gracious response to the world, but a tendency one can find in moral philosophy.

When we engage in moral philosophy, or philosophical ethics more generally, we often assume what already amounts to something like a prescriptive mood. We look for what sorts of things a person ought to do, and how they ought to be. How, as legislators, ought we legislate? How, as members of a moral community, do we think our fellow members ought to conduct themselves? And even when we are less obviously prescriptive, our thoughts about one another can nonetheless be informed by considerations that encourage a similar stance. When considering what the virtues are, for example, we are often asked to wonder what sorts of qualities we would want our *children* to have. This gets us in the prescriptive mood once again, given that it is a guiding, if not constitutive goal of parents to raise their children to be *good*, for their children's own sake as well as for the sake of those who form one's community.⁴⁹

The sort of attitude that I am discussing here is one that can get easily overshadowed by those we take on when we reflect from these more prescriptive perspectives. One response to this tendency to moralize in moral philosophy is to think that we should do away with evaluative evaluation, entirely. But I think that this would be a mistake – it would encourage us to be less sensitive to important features about one another and ourselves that are good to be aware of, even when we aren't in the business of moralizing *about* them.

An alternative perspective that we often occupy, and which better encourages the attitude I've described, is perhaps best illustrated – for several reasons – by turning our attention away from reflecting on the qualities we'd want our fellow citizens, community members, or children to have, toward reflection about what sorts of characters we would want our *friends* to have. Our

⁴⁹ It might be this that makes it difficult for Jack's father to experience grace toward him: he occupies a particular sort of role that Glory does not – as his parent it was his responsibility to raise a *good* child.

friends are beings whom we love, whom we love as our equals, but whom we need to view in a way that puts limits on what genuinely well-intentioned actions and thoughts done and had for their sake, or for the sake of others, could be appropriate.

In a recent article, Benjamin Bagley has made a point that is similar to what I am suggesting here. In objecting to the idea that we can understand love as just a blanket desire for another's wellbeing, he writes,

I care about my close friends, and would do a lot for them if they asked me (and in some cases even if they didn't). But there are a lot of helpful things I could do for my friends that I feel absolutely no desire to do: their laundry, for instance, or their grocery shopping. It's not that such desires are overridden by others, or that my friends would find it off-putting if I acted on them... It's rather these sorts of things just aren't what friendship is about.⁵⁰

The basic thought here – that friendship just isn't centrally about helping one's friends do well – is one that I think should be extended to the *characters* of one's friends, as well. I'd neither expect nor want my friends to do my laundry or buy my groceries, and when it comes to the bad aspects of my character – though I prefer that they be well aware of those aspects – they would be right to set limits on their willingness to do my chores as well.

This suggests that it is important for us to supplement, or limit, ways of thinking and evaluating that are more imperatively-minded, at least when we are just in the everyday business of having certain sorts of relationships with one another. We need to allow ourselves and others the space from which we might better be able to *appreciate*, and not just monitor, hope to improve upon, or even police the characters of others, even when genuinely for their sake. This need not indicate a *failure* in one's love – instead, it may be a manifestation of it.

⁵⁰ Bagley, "Loving Someone in Particular" p. 482. Bagley's article discusses the puzzle about love that I have not tried to handle here: that of how we can be justified in loving one person in particular.

IV. CONCLUSION

I've argued that there are virtues to knowingly loving the bad, including at least one way of understanding the virtue of graciousness. But one might justifiably wonder whether there are limits to what could count as an exhibition of these virtues, given just how bad some people can be. Are there examples of people or qualities of character that are just so terrible that love of those people or those qualities would not be underwritten by virtue, but something else? It is instead a question about the limits of grace, itself. Perhaps Glory's love for Jack is something commendable or admirable, but what if Jack were considerably *worse* than what he is actually like? Would we not begin to wonder whether she is, after all, projecting something where there is nothing, or being insensitive to just how bad he is?

Related to this question, there is of course the important practical question of *what to do* with such people. If Jack's "streak of malice" were, for example, directed toward Glory, presumably this would give her a salient reason to give up continuing in her relationship with him. The problem becomes more serious when we think of the ways in which people are seriously harmed and wronged by those whom they love. Does it follow from concluding that graciousness is a virtue that it is a good and sane thing for people to put up with or endure those who seriously abuse them, or others?

Two minor suggestions is that graciousness, like all virtues, may have vices of deficiency and excess, and must be balanced with others, including prudence and self-respect. But these are rather flimsy points: it is a good question how far the limits of grace itself might go. I don't here have an answer. But in wondering about this, we might, as a starting point, return again to the example of Helen Prejean, and her love of Patrick Sonnier. This is, undoubtedly, a difficult sort of case to use as a guide, given what Prejean knows about Sonnier, and given that Prejean is working within a theological background that is not widely shared. We may be tempted to see

her love for Sonnier simply as the sort of love that doesn't focus on *his particular character*, but rather, on him as simply the human being who happened to cross her path and who is equally deserving of a kind of universal love.

But we might try to imagine that instead, she loves Sonnier in a way that is informed by what he is like – having met him, and having found that though he is bad, his badness is nonetheless a kind of human badness. And in considering *our* reactions to her love, we might remind ourselves about what graciousness is like, and what its connections to other virtues are. We might consider what exactly it is about ourselves that makes it difficult for us to love such a person, and look again. It may be helpful to remind ourselves that love is not inconsistent with thinking that the beloved deserves to be *punished* for the bad things that he does; and that reactive attitudes of anger and blame are appropriate and warranted with those we may in fact be closest to. Indeed, when I reflect on the aspects of my own character toward which my friends exhibit their graciousness, I find it most familiar that their reactions are a mix of both affection, *and* exasperation. There is no inconsistency or incoherence in this.

And it also coherent to love a person even if one must extricate oneself from an ongoing relationship from her, or while allowing her to extricate herself. Perhaps this is why – as I suggested above – grace is perhaps most easily experienced from the perspective of friendship, rather than other relationships. These other relationships, such as the relationship between a parent and a child, or the relationship between romantic lovers or spouses, are characterized by a kind of active *involvement* with that need not define even the very excellent friendships. These other sorts of relationships involve moving forward together, and certain vices and bad tendencies are simply inimical to that kind of development. You cannot move forward with another if, for example, she is emotionally closed off, or because she consistently harms herself,

or others, or you. Of course, sometimes our friends are like our family, in which case the tension arises; but friends may love another without always having the same sort of involvement that other kinds of relationship call for.

Given the kind of answers that I have provided to these worries, one might begin to wonder: what is the *good* of this form of love if one could feel it without having to *do* anything for the beloved? Why should one bother with it if it does not even demand that one continue on in an ongoing relationship with him, or that one must be centrally concerned with his wellbeing? Indeed, Robinson's novel concludes with Jack leaving home once again, unable to involve himself in the life of his family any longer, thinner and wearier than before. From a certain perspective, one might think that Glory's love *made no difference*.

In light of these questions, perhaps it would be best to return to the larger discussion in which Murdoch provides her example of M and D, in the first place. There, Murdoch is not simply illustrating a particular kind of moral achievement. She is responding to a certain paradigm of ethical thought that she believes we should resist, or at least, place certain limits on. Such a paradigm, in focusing exclusively on action and public states of affairs, cannot accommodate the fact that M's shift in perspective toward D is an important one, and that it is important even if it results in no difference in her outward behaviour toward D, and even if there is no way – either through death, or distance – for their relationship to continue.⁵¹ A shift in perspective and attitude toward another and the world may be entirely private, but no less profound because of this.

⁵¹ "Let us assume for purposes of the example that the mother, who is a very 'correct' person, behaves beautifully to the girl throughout, not allowing her real opinion to appear in any way. We might underline this aspect of the example by supposing that the young couple have emigrated or that D is now dead: the point being to ensure that whatever is in question as *happening* happens entirely in M's mind" (p. 17).

The Undesirable and the Adesirable

I. INTRODUCTION

In the last chapter, I argued that those who can knowingly love the bad – in particular, the badness of human beings – might be exhibiting the virtues of, among others, graciousness and a form of self-esteem. But one might wonder whether I’ve really shown that conclusion, because it seems to depend on a rejection of the thought that our desires, loves, and other positive attitudes can correctly aim at something that isn’t good at all. Surely there must be *something* good about an object that makes sense of our loving or desiring it – otherwise it is hard to see what could be different between these sorts of attitudes (which I’ve suggested can be intelligent responses to the qualities of an object), and a blind or compulsive attraction.

Indeed, persistent throughout philosophical history is the claim that we necessarily desire or intend in light of some goodness, apparent to us, in what we desire or intend. I will here refer to this as *the guise of the good* thesis. Not only is some version of the claim recognizable in the thought of figures such as Socrates, Plato, Aristotle, Spinoza, Aquinas, Hobbes, Kant, and Sartre, it has recently re-emerged in contemporary philosophy, defended in some form by, among others, David Gauthier, Dennis Stampe, Warren Quinn, Barbara Herman, Christine Korsgaard, Thomas Scanlon, Sergio Tenenbaum, and Joseph Raz.⁵²

⁵² David Gauthier, *Morals by Agreement*, (Oxford: Oxford University Press, 1986). Dennis Stampe, “The Authority of Desire,” *The Philosophical Review* 96 (1987), 335–81. Warren Quinn. “Putting Rationality in Its Place,” in

A glance at these lists and one may be struck by the ubiquity of the claim across philosophical outlooks that are otherwise vastly divergent. This might suggest to some that the thesis is undeniable – indeed, Donald Davidson once wrote that it has “an air of self-evidence”.⁵³ But, as with any sweeping philosophical claim, there are others who are just as certain of its falsity. Michael Stocker, for example, reports that the seeming pervasiveness of the claim that we only desire the believed good “affords [him] no pleasure, since that view... is clearly and simply false”.⁵⁴ The rest of us may find ourselves wavering somewhere in between – perhaps sensing that there is something about the thesis that seems right, while also having some inkling that it seems altogether too neat.

Much hangs on the issue, outside of showing the conceptual and psychological possibility of the sort of attitudes I described in my first chapter. The thesis sets limits on how to understand the phenomenon of *akrasia*, as well as cases in which people seem, perhaps perversely, motivated to do something *because* it would be bad to do. Moreover, many of those who endorse

Warren Quinn (ed.), *Morality and Action*, (New York: Cambridge University Press, 1993). Barbara Herman, “Making Room for Character” in *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, Stephen Engstrom & Jennifer Whiting (eds.) (Cambridge: Cambridge University Press, 1996). Christine Korsgaard, *The Sources of Normativity*, (New York: Cambridge University Press, 1996). T.M. Scanlon, *What We Owe To Each Other*, (Cambridge, MA: Harvard University Press, 1998). Sergio Tenenbaum, *Appearances of the Good: An Essay on the Nature of Practical Reason*, (New York: Cambridge University Press, 2007). *Desire, Practical Reason, and the Good*, (Oxford: Oxford University Press, 2010), pp. 65–81. Joseph Raz, “The Guise of the Good,” in Tenenbaum (2010).

⁵³ Donald Davidson, “How is Weakness of Will Possible?” in *Essays on Actions and Events*. Davidson himself took it to be self-evident that we desire the believed good, and that this commitment is part of a standard denial of the possibility of *akrasia*.

⁵⁴ Michael Stocker, “Desiring the Bad: An Essay in Moral Psychology,” *The Journal of Philosophy* 76, 12, (1979) 738–53, 739. Gary Watson, Peter Railton, and David Velleman, either explicitly argue against the thesis, or have developed views about the nature of human motivation that entail its denial. Gary Watson, “Free Agency,” *The Journal of Philosophy* 72, 8 (1975), 205–20. Peter Railton, “On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action,” in G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason*. (New York: Oxford University Press, 1997), 53–79. Velleman, David, “The Guise of the Good,” *Noûs* 26, 1, (1992) 3–26.

the thesis do so because it serves as a premise in certain attractive theoretical projects.⁵⁵ And as I will discuss, it gives rise to a certain way of understanding the justificatory nature of our motivational states, and our relationship to goodness as possessors of those states.

But the terrain forward can be rough going. The seeming ubiquity of the thesis is partly explained by the fact that there are many ways to interpret it, each differing in content and plausibility. There is the growing sense that the trading of examples and counterexamples has been dissatisfying, or that the debate is best deflated. And there are methodological divisions that need to be kept in mind: at least some of those who defend the thesis see it as one piece of a Neo-Aristotelian picture replete with natural goodness and teleology that many who deny it cannot stomach.

In what follows, I shall outline a conciliatory alternative. Part of what is unsatisfying about a straightforward denial of the thesis is the nagging sense that cases in which there is *nothing* apparently good in what one purports to desire strike as involving *not* desire and action, but inclination, compulsion, or mechanism, instead. Proponents of the guise of the good thesis account for this difference by maintaining that the mental states that are intelligible as desires are sensitive to considerations about the desirability or undesirability of their objects, understood in a normative sense. Defenders of the guise of the good thesis are right to claim this but, as I'll argue, the evaluative properties that they are sensitive to needn't be identical to goodness and

⁵⁵ Raz writes, and Tenenbaum concurs, that the guise of the good thesis is, "the keystone, keeping in place and bridging the theory of value, the theory of normativity and rationality, and the understanding of an intentional action" (2010, 134). It has also been used in some arguments to show the rationally obligatory nature of *moral* demands. And one might be attracted to the thesis' potential role in establishing what some call the "unity of the normative". For example, if desires "aim" at goodness in the same sense that beliefs "aim" at truth, then perhaps we will be able to reach some general normative principles that govern both beliefs and desires, and epistemology and ethics can be brought together under a broad "meta-normative" project. For a recent paper that relies on the parallelism of desire and belief as evidence for the truth of the guise of the good thesis, see Karl Schafer, "Perception and the Rational Force of Desire." I take the phrase "the unity of the normative" from "The Unity of the Normative" by T.M. Scanlon, who there expresses his hope that "the normative domain [consists] simply of truths about reasons, including reasons for actions and reasons for belief and other attitudes" (p. 443).

badness. Indeed, assuming this identity threatens to overlook the dimensions of an important, ethical evaluation – one to which our interpersonal reactions and attitudes are sensitive.

We must distinguish between two substantive senses of *desirable*. One may be synonymous with the good or the apparently good, but the other simply refers to those things that human beings, given their nature, are *able* to desire. In drawing this distinction, we should resist the thought, encouraged by G.E. Moore, that there are only two senses of “desirable” available to us: one that refers to the good, and another that is purely descriptive, naming anything that we happen to find actual human beings striving for.⁵⁶ There are normative constraints on what human beings can desire, constraints that are imposed by human nature – just as Neo-Aristotelians suggest. But those constraints are not identical to considerations about apparent goodness. Our desires, I will argue, are under the guise of the *naturally attractive*, not the good. Likewise, the “undesirable” (that is, the bad) is not identical to what I will here call the *adesirable*: those things that cannot be the objects of human desire, at all.

II. WHAT IS THE GUISE OF THE GOOD THESIS?

Though many accept something like the basic thought that our intentions and desires are under a guise of goodness, it can be difficult to isolate a particular claim that is both plausible and substantive. Indeed, one might suspect that the alternative thesis that I will argue for – that human desires are under the guise of the naturally attractive – is either veering on truistic emptiness or is simply a slight variation of the thesis if distinct from it at all.

The thesis that desires are under the guise of the naturally attractive is not empty, and is importantly different from the guise of the good thesis. In order to demonstrate these points, as well as get a sense of what it could mean for our desires to be “under a guise” at all, I will begin

⁵⁶ G. E. Moore, *Principia Ethica*, p. 67.

by articulating and motivating the guise of the good thesis. In light of the reasons that speak in its favour, as well as the reasons for finding a certain common way of rejecting it unsatisfying, it will be easier to see what my alternative position amounts to.

An initial, rough statement of the guise of the good thesis is this: that when a human being desires something, or performs an intentional action, she must be doing so in light of something good she takes there to be in what she desires or intends to do. But this is very rough, indeed. What sort of phenomena are we picking out, and what are we contrasting them to, when we isolate “desire” and “intentional action”? In what sense “must” a human being take there to be something good in what she desires? In what sense must she “take” there to be something good in what she desires? And in what sense of “goodness” must she take something to be?

While different variants of the thesis will provide different answers to these questions, I will provide those that keep in mind a particular, central motivation for the thesis. This is the thought that we can only understand an agent’s motivations and behaviour as intelligible desires and actions as opposed to tics or compulsions or mechanistic operations, should there be some good, apparent to her in the object of her desire and intention, by which she is motivated. The thesis is no mere standard by which it would be appropriate to *assess* or *criticize* a person’s desires or actions. Rather, in order to understand a person’s motivation *as a desire* and the resulting behaviour *as an action*, there must be some aspect of what she desires or intends that strikes her as good. Anscombe memorably illustrates this line of thought, writing:

It will be instructive to anyone who thinks [that anything can be the object of a desire] to approach someone and say: ‘I want a saucer of mud’ or ‘I want a twig of mountain ash’. He is likely to be asked what for; to which let him reply that he does not want it for anything, he just wants it...⁵⁷

⁵⁷ Anscombe, *Intention*, p. 70. It is a good question whether Anscombe actually endorses the guise of the good thesis in *Intention*, as she also states that intentional actions can be done “for no reason”, and so (we can safely assume) for no reason that consists in some apparent good (25). For my purposes, it is enough that she is canonically read as endorsing the thesis.

She then suggests that as long as we are taking this man seriously, there are two ways for us to go on. One is to determine what it is that this man takes to be good, or desirable, about having a saucer of mud, or twig of mountain ash. But if we can find no such “desirability condition,” we will have to defer to our second option and conclude that this man is, as Anscombe puts it, “a babbling loon”.⁵⁸

One central lesson we are meant to draw from her discussion is that our desires are not simply brute and blind; rather, they are sensitive to considerations for or against desiring things. And in particular, Anscombe suggests, they are sensitive to considerations about the *goodness* of what is wanted or intended. When a person’s desire seems utterly indifferent to such considerations – when the man who wants a saucer of mud simultaneously sees *nothing* good about having it, he “just wants it” – we encounter a certain kind of difficulty in understanding him. We may get the sense that to think of the man as having a *desire* isn’t right, perhaps concluding instead that he is experiencing a blind urge. A person suffering from OCD, for example, who washes her hands seven times before leaving home, need not see this as something *good* to do – nor would we think of her motivations as desires.⁵⁹

⁵⁸ *ibid*, p. 70.

⁵⁹ One could motivate a similar argument by assuming a more Kantian approach to the mind. For example, one might start with the basic thought that desires are in some sense necessary for action, but in order to understand an agent as acting *freely*, or (to avoid the connotations of that terminology) without being *compelled* by her desires, she must also have the capacity to act only on those desires that are also for something that she reflectively endorses, or takes to be *good* in some sense. Should a person fail to “act” under such a guise, this “action” is really, again, best understood as a kind of compulsion or determination by her desires or inclinations. For Andrews Reath, Christine Korsgaard, and Alan Gewirth, the guise of the good thesis is one step along the way to showing that moral demands are rationally overriding, as acting against them involves acting against commitments one already endorses (simply in virtue of being a rational agent). This is one way to understand how Kant argued for such a conclusion, but Thomas Hill has argued that while Kant certainly endorses the claim that one must act under the guise of some personal good, it is an open question whether the stronger version of the claim defended by Reath and Korsgaard (that the object must be taken to be in a sense, objectively good) is true. There is a way to apply the arguments I will give here to this neo-Kantian defense of the guise of the good thesis, but I will not do so explicitly. Gewirth, *Reason and Morality*, (Chicago: Chicago University Press, 1980). Korsgaard, “Kant’s Formula of Humanity”, *Kant-Studien*

This argument in favour of the thesis relies on a rough, intuitive sense of a distinction between those motivations that are intelligible as desires and those that are not. It is worth considering why this distinction is important for our assessments of one another more generally. This will help us maintain a clearer sense of what sort of “intelligibility” here is at stake. As Anscombe elaborates on this point, when we encounter this kind of unintelligibility, we are not “excluding a form of words from the language; we are saying ‘we cannot understand such a man’”.⁶⁰ Moreover, we should notice that encountering this form of obscurity in another person brings with it a shift in our reactions towards him. Rather than criticizing or correcting him for his purported “desire”, we may find ourselves withdrawing such reactions. As Peter Strawson puts the point, “Seeing an agent [as one whose picture of the world is an insane delusion; or as one whose behaviour, or a part of whose behaviour is unintelligible to us]... tends to inhibit ordinary interpersonal attitudes in general”.⁶¹

These ordinary interpersonal attitudes, as Strawson elaborates, include but are not at all limited to: resentment and indignation, gratitude and forgiveness, certain forms of love or affection, and whatever other unnamed emotional responses we have towards a person when we view him as *participant* in a human relationship and life, rather than as an object of management, treatment, training, or scientific observation.

Where we might criticize a person’s desire to do something cruel, or praise someone’s desire to do something kind – using standards that are *relevant* to assessing one’s desires, but not constitutive of them – we withdraw either sort of reaction when his motivations and behaviour

77 (1-4), 1986. Reath, “Did Kant Hold that Rational Volition is Sub Ratione Boni?” in Robert Johnson and Mark Timmons, eds. *Reason, Value and Respect: Kantian Themes from the Philosophy of Thomas E. Hill, Jr.* (Oxford: Oxford University Press, *forthcoming*).

⁶⁰ *Intention*, p. 27

⁶¹ Strawson, “Freedom and Resentment”, pp. 16-17.

are unintelligible. Rather than criticize or correct him, we are instead inclined to manage him, treat him, study him, or simply keep our distance. We take what Strawson calls *the objective stance* towards him – at least in regards to this aspect of his psychology. The emotional reactions that we have from this stance, if any, will include “non-reactive” ones like pity, fear, and even forms of annoyance – but not indignation or resentment. And if we cannot restore a certain level of intelligibility to his mind and behaviour, it is difficult to so much as raise the question of whether he is *justified* in doing what he is doing, or whether he is appropriately subject to other sorts of thick ethical evaluations. Such reactions and evaluations have as a basic background supposition the intelligibility of the object of our assessment. Intelligibility of this sort, then, is not a “normative” category in the sense that it supplies agents with prescriptive reasons to desire or act in intelligible ways; but it is “normative” in (1) the sense that it is important, if not indispensable, for our ethical reactions to one another in light of our motivations and actions, and (2) that what it is sensitive to (at least according to the guise of the good thesis) are considerations that are themselves evaluative.

Given this basic motivation for the guise of the good thesis, we can now begin to sharpen the very rough statement I began with. The “must”, as we have seen, is one that refers to a constitutive norm that governs which of our mental states *count* as desires and which do not. Moreover, we can disentangle the thesis from others that are sometimes discussed in tandem. We can isolate it from a similar, but distinct claim – a version of motivational or judgment internalism – that apparent goodness *always* attracts. And a defender of the guise of the good thesis need not be committed to the claims that she takes the object as good *overall*, or that a person cannot knowingly or willingly do or desire something that is on balance worse than his other options. Rather, she can allow for certain forms of *akrasia*, simply adding that when a

person acts against his judgment about what would be all-things-considered best, it is because he nonetheless still desires, chooses, or acts in light of *something* good that he sees in the worse option.

IN WHAT SENSE MUST THE AGENT *TAKE* THE OBJECT AS GOOD?

Another question we were left with is in what sense must the agent *take* there to be something good in what she purportedly desires, or intends to do? Different variants of the view answer this differently. One might think, for example, that the agent must have a *belief* that the object is good in some way.⁶² But the thesis is more plausible without this particular commitment, for at least two reasons. The first is the familiar worry that this would be implausibly demanding. Presumably there are some beings that have desires and perform actions without possessing the cognitive and conceptual capacities needed for forming evaluative beliefs: non-human animals and very young children come to mind. The second is that this version of the thesis seems to distort, or conflate, ordinary cases of desire in ways that proponents of the thesis should recognize. A person might desire a saucer of mud because, for example, his friend has dared him to eat it, without at the same time possessing the *belief* that this is *good* to do this thing. And somebody who did this because he believed that it was *good* to do so, or good to act on dares, would have a different sort of character or evaluative outlook on

⁶² David Velleman (1992) argues against several different versions of the thesis, first focusing on a version that relies upon the agent's capacity to use evaluative concepts, before then moving to a version that attempts to capture the thesis in terms of direction of fit. His arguments against both versions rely upon defenders of the thesis simultaneously maintaining certain commitments that I will not hold them to: in particular that justifications of actions takes a specific, inferential or deductive form. As Velleman shows, these further commitments lead proponents of the thesis into difficulty, but his discussion may leave us wondering whether some essential core of the thesis – stripped of certain theoretical ambitions and without commitment to that particular view of justification – is nonetheless right.

what it is to act on a dare than the first. Such a person would be a bit more fastidious, or neurotic, or just plain weirder than the first person; and we should be able to mark that difference.⁶³

Given these worries, we should interpret the relevant “guise”, or “taking” as a quasi-perception, or appearance, or construal of goodness, rather than as a conceptually laden evaluative belief or judgment.⁶⁴ Some might worry that this is too opaque: what does it mean to *construe* or *quasi-perceive* something as good, if this means something more than finding it attractive? There are a number of ways we could spell this out, each depending on what substantive conception of goodness a defender of the claim has in mind. Rather than exploring those options, I suggest that we interpret this commitment in the following manner that is neutral between them: that should a human being desire something or perform an action that *in fact* has nothing good about it, or which may be positively bad, we should conclude that she is making a kind of mistake about the evaluative qualities of the object, or action.⁶⁵ This is because, according to the thesis, the *target* or *aim* of intelligible human desires and intentional action is *something* good. Moreover, this perception or construal of goodness is what gives rise to one’s

⁶³ In “Desiring the Bad Under the Guise of the Good”, Jennifer Hawkins makes a similar point, using a case of a child who engages in a kind of ritual when she eats her food – switching between each individual item between bites, so that her lunch “comes out even”. Hawkins writes, “To say that she *values* the ritual, or has some belief about its worth, would seriously distort the case” (249). I do not think that a defender of this version of the thesis would have to commit to the thought that the child *values* the ritual, but I do think that it already distorts the case to suggest that the child has a *belief* that it is good. Beyond worries about whether she has the conceptual resources to form this sort of belief, she would, if nothing else, be a little *stranger* than her counterpart who simply engages in this simple ritual.

⁶⁴ This is consistent with a general thesis in the philosophy of emotion that our emotions are not composed of beliefs about the evaluative landscape, but rather, *construals* of goodness and badness. It is also what particular defenders of the thesis have explicitly endorsed. See for example Stampe (1987), Tenenbaum (2007) and Graham Oddie’s *Value, Reality, and Desire*.

⁶⁵ This makes it especially difficult to understand the guise of the good thesis as a substantive claim when maintained by those who would also deny that evaluative judgments could be mistaken, perhaps because of a commitment to a radical sort of subjectivism or anti-realism about evaluative facts. But perhaps this is just as things could be: as others have noted, such views would be committed to a trivial version of the thesis because they maintain that desires just are claims about goodness. Indeed, as Tenenbaum argues, one way to deny that there is a substantive guise of the good thesis is to endorse subjectivism about value.

desire, precisely because according to this thesis, it is what desires are themselves sensitive to. While this answer does not alone address the difficult question of what it is for an agent to *construe* something in a certain way, and how she could do so without making use of concepts, it nonetheless allows us to see how the guise of the good thesis provides both constitutive standards on desire, as well as a standard of correctness.

Moreover, these two elaborations help us get a firmer grip on what hangs on the truth or falsity of the guise of the good thesis, and also capture one important historical motivation for it: the thought that when one desires or intends something that is in fact bad, it is the result of a hazy, confused, or otherwise less than ideal perceptual or cognitive grip on the facts of one's situation.⁶⁶ And this provides us with some resources to understand what is going on in certain cases of *akrasia*: I might judge, or even know in some sense that I ought not attack my squash opponent after an ignominious defeat, but in the grips of my frustration and embarrassment, I cannot but suffer from the illusion that doing so would be something good to do.

IN WHAT SENSE MUST THE AGENT TAKE THE OBJECT AS *GOOD*?

The final question we are left with is in what sense must the agent construe the object of her desire as *good*? While again, there are different ways in which this aspect of the thesis could be spelled out, some proposals are more plausible than others. For example, we should reject theses which propose some singular, substantive good that the agent is acting in light of, such as her own wellbeing, or moral goodness. Instead, the thesis should be paired with the most plausible, pluralistic understanding of value we have available to us.

⁶⁶ For example, Socrates' denial of the possibility of *akrasia*, and Aristotle's claim that *akrasia* is the result of a person's having a deficient sort of knowledge of the badness of what she is doing, both reflect this basic idea.

However, this flexibility must be constrained once we recall what sort of work desirability conditions are meant to do. Notice that it doesn't help our puzzlement, for example, to learn that Anscombe's man perceives or construes the *having of mud* or the *having of saucers of things* as intrinsically good. A successful desirability condition does not just involve the thought that he views the desired object or action as good in *any* empty sense: it must reference some aspect of the object or action that a person could plausibly take to be good.⁶⁷ This condition of plausibility helps us retain the guise of the good thesis as a substantive claim, disallowing for trivial or empty citations of goodness, and so, trivial or empty versions of the thesis.

Let me summarize the key points from these last three sub-sections before considering a common and illuminating objection to the thesis. The central motivation for the guise of the good thesis is that the difference between desires and intentional actions on the one hand, and behaviour like tics or compulsions (and the mental states that give rise to such things) on the other, is that the former are "under a guise of goodness", while the latter are not. This "guise", I've argued, is best understood as a quasi-perception, or a construal of goodness, rather than a belief or judgment, where the apparent goodness of the object is what gives rise to a person's desire. And importantly, it follows from this commitment that should a person desire something that is *in fact* not good, she is thereby making a mistake akin to the mistake one makes when suffering from a perceptual illusion. Even if one *knows* that the Müller-Lyer lines are even, for example, one's perception is still misleading – so one can *know* (in the same sense) that what one wants is not good, but that one desires it intelligibly is a sign that one's "perception" of it is misleading. Finally, the sort of goodness that the agent sees must be something that it is *plausible* for a person to find good. Without being able to determine such a desirability condition, we will

⁶⁷ This is why, as Anscombe suggests, we can understand the desires of a Nazi officer. We do not think that genocide is plausibly good, but we do think that it is plausible that a human being *would think it good*, mistakenly.

– as Anscombe suggests – have to conclude that the supposed agent is in this respect unintelligible, where importantly, such a verdict brings with it a shift in the kinds of attitudes that it is fitting to have towards him, at least in regards to this particular aspect of his psychology.

III. AN ILLUMINATING OBJECTION

With this elaboration of the thesis in mind, we are in a better position to see why a common denial of it should strike us as dissatisfying. This particular denial emerges from an allegiance to an alternative theory of human desire and action. Perhaps *the* standard alternative is one that is best labeled (though not for reasons of historical accuracy), *Humean*.⁶⁸ According to the Humean, there exist two basic mental states with two different functions: one representational (Humean “beliefs”), and the other motivational (Humean “desires”). Humean “desires” are necessary for the performance of an action, and although they may have certain formal constraints (such as a direction of fit, and that they take as their objects states of affairs), they need not specify that what is to be brought about has qualities that the agent sees as *good*. Indeed, such states *lack* representational content, and so would *a fortiori* lack that particular representational content. While Humean “beliefs” might contain this content, they are motivationally inert and need to be paired with a desire in order for an agent to perform an action.

Moreover, no such guise is needed for a person’s desire to be intelligible as such. For example, take a discussion from Michael Smith’s original defense of the theory. There, he suggests that knowing of a man who has a “yen” to drink a can of paint that he has always had

⁶⁸ There are, of course, various versions of “Humeanism” about action, mental states, and practical reasoning, etc. and I do not mean to conflate them here. This is instead meant to serve as an illustration of a general sort of objection that one might have to the guise of the good thesis, rather than an argument against any particular version of Humeanism.

this yen is enough to lend him a “partial justification” for drinking the paint.⁶⁹ For Smith, this “partial justification” lies in the agent’s having a *motivating reason* that, unlike a *normative reason*, explains the action in terms of the agent’s possession of a motivational state, rather than justifying the action according to some independent requirement.⁷⁰

When considering this example in abstraction from any background detail, we may find appealing the thought that this episode of paint drinking really is no different from other actions, given that it shares with them a certain formal structure. An agent desired or had some “pro-attitude” toward *x* and did whatever it was necessary to have or bring *x* about. But it is difficult to find this a satisfying response to the particular challenge raised by defenders of the guise of the good thesis. If we do not find a person’s drinking of a can of paint intelligible, learning that he has always had this “yen” to do so doesn’t provide us any illumination when he finally takes a gulp. Likewise, that a person has a motivational disposition to eat clumps of hair or scratch herself until she bleeds doesn’t help us to understand her behaviour.⁷¹ The “partial justification”

⁶⁹ This example is originally given by Davidson in “Actions, Reasons, and Causes”. He writes, “...a man may all his life have a yen, say, to drink a can of paint, without ever, even at the moment he yields, believing it would be worth doing” (p. 4).

⁷⁰ Michael Smith, “The Humean Theory of Motivation”. There, Smith seems to endorse some version of the guise of the good thesis, as he writes that the man is justified “from a perspective that *assigns value* to the drinking of a can of paint” which he occupies “only to the extent that he has a yen to drink [it]” (p. 39, my emphasis). I think, however, that this is all a bit too obscure to amount to a clear endorsement of the guise of the good thesis, and in any case, the Humean Theory need not be committed to it. For a more recent discussion of the Humean Theory, see Neil Sinhababu, “The Humean Theory of Motivation Reformulated and Defended”.

⁷¹ Kieran Setiya offers an alternative theory of action that denies the guise of the good thesis, but seems to run into the same problem as the Humean theory. He argues that we do not need to act under the guise of a “normative” reason, but only an explanatory one. In response to a case of Thomas Nagel’s (of a person trying to put a coin into a pencil sharpener in order to get a drink from a nearby vending machine) Setiya suggests that learning that this person has been *conditioned* to have this motivational disposition *does* make his behaviour intelligible to us as an action, as we now have an explanation for his behaviour (65-67). He then suggests that such cases are “rare”, but not impossible. But his assessment of the case repeats the unsatisfying point that makes the guise of the good thesis compelling. I agree with Setiya that we should not think that these cases are impossible, but that we face a kind of difficulty seeing this behaviour as an exercise of agency. Such cases can be made intelligible with enough of a background story – but importantly, this background story is needed and furthermore, is precisely what defenders of the guise of the good thesis will argue must reveal a desirability condition if this person is to become intelligible to us. The mere motivational disposition, even when we are aware it has been brought about by conditioning or

that she has to satisfy these urges, whatever it amounts to, is not one that is relevant to addressing the concern that defenders of the thesis raise. In other words, the notion of “desire” used in such views is so thin that we are left without a way of distinguishing – of all our motivationally potent mental states – which are expressions of the sort of human agency that can merit our interpersonal attitudes and our ethical assessments, and which can do nothing of the sort. If we want to retain this thinner notion of desire, we will have to supplement our psychological theories with the desires that defenders of the thesis are trying to account for – in which case the thesis re-emerges as a potential analysis of this important subset of our motivational attitudes and behaviour.⁷² In short, the Humean theory of motivation erases the distinction that the guise of the good thesis is trying to capture.

Perhaps there is a way to mark this distinction that is consistent with the basic Humean framework. But there is a dilemma that defenders of the guise of the good thesis should press on attempts at doing so. There are, of course, Humean views that rely on further formal constraints: for example, we might try to distinguish these bizarre cases of motivation from intelligible cases of desire by relying on coherence between an agent’s desires, or a kind of second-order endorsement. And Humeans sometimes claim that agents whose psychologies satisfy particular

training, fails to do the work – it may even deepen our sense that this person is operating mechanistically. Setiya, *Reasons Without Rationalism*, (Princeton: Princeton University Press, 2007).

⁷²As an illustration of this dialectic, consider James Lenman’s recent criticism of Anscombe’s discussion. Lenman grants that there is something unintelligible about a man who wants a saucer of mud, but argues that this only suggests the existence of a formal constraint on desire: “For bare intelligibility we don’t need a desirability characteristic, just an explicit specification of propositional content”. Tenenbaum echoes this point: “...a saucer of mud is neither an action nor a state of affairs. The fact that one cannot want a saucer of mud is arguably no more surprising than the fact that one can not believe a saucer of mud; ‘saucer of mud’ is simply not a proper complement to ‘wants (to/that)’ or ‘believes (that)’”. But the sense of intelligibility that I am focusing on here won’t be provided in these cases by merely restating their content in propositional form. We can specify the “desire” of the compulsive hand washer in this way – *I want that I wash my hands seven more times*, or of the mud-wanter – *I want to eat a saucer of mud*, but this does not help us understand this motivated behaviour as a desire in the relevant sense. Lenman, “The Saucer of Mud, the Kudzu Vine and the Uxorious Cheetah: Against Neo-Aristotelian Naturalism in Metaethics,” *European Journal of Analytic Philosophy* 1 (2); 37–50. Tenenbaum, “Guise of the Good”, *The International Encyclopedia of Ethics*, First Edition, Hugh LaFollette (ed), (Blackwell Publishing Ltd., 2013).

Humean-friendly standards of rationality, including full non-normative information, will turn out to *lack* these sorts of bizarre motivational states.⁷³ But either these added constraints will include something that amounts to an endorsement of the guise of the good thesis, or they will not do the trick. For example, learning that the paint-drinker's "yen" is one that he also identifies with, or desires to be effective in his actions, renders him even more bizarre, not less so. If paint-drinking (or wanting a saucer of mud or twig of mountain ash) is not the right sort of object of a desire in this sense (so a defender of the guise of the good thesis will point out) it is not because it fails to cohere or conform with an agent's other interests or because he doesn't endorse it all the way down: it's because there is something *about* paint-drinking that cannot be fully captured by the non-normative facts about it that makes it so. It's because there's *nothing plausibly good* in doing this sort of thing. The worry, in short, is not that these alternative views cannot put constraints on what is desirable and what is not – it is that the constraints they place are either not of the right kind, or will be an endorsement of the very thesis that is being resisted.

IV. THE HUMAN BEING & THE HUMAN GOOD

One explanation for why this last sort of view fails to make contact with the guise of the good thesis is because of a large methodological difference between views of human mental states that seek to reduce and naturalize, and those that operate by introducing notions that are already, in a certain sense, normative – or at least, difficult to square with a reductive, naturalistic

⁷³ In defense of the Humean theory, Derek Baker writes (in considering the importance of intelligibility): "Since agents will typically (perhaps necessarily) meet some minimal threshold of rationality, most of their choices will conform to their background beliefs about what they ought to do." He also suggests that, "since the Humean can hold that desires have a (nonrepresentational) phenomenology, they can explain any default intelligibility of our actions in terms of feelings of excitement, pleasure, and distress that the potential satisfaction and distress of our desires brings" (22). But this point suggests to me that we must be trying to capture different sense of intelligibility. I very well know that a person with a compulsion to eat her hair may be distressed if she is prevented from doing so, but this does very little (if anything) to achieve the sort of intelligibility under discussion here. Baker, "The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire," *Journal of Ethics and Social Philosophy*, 8, 2 (2014).

enterprise. Indeed, the guise of the good thesis is perhaps at its strongest when underpinned by a philosophical outlook that appeals to a robust conception of human nature, the human good, and a teleological model of explanation that renders our activity intelligible only when understood in light of those concepts. I will refer to this as the *neo-Aristotelian* defense of the guise of the good thesis. Not all defenders of the guise of the good thesis also maintain these neo-Aristotelian commitments, but discussing the thesis with such commitments in mind may illuminate the motivations that others may have, as well as clarify certain aspects of the thesis that can be otherwise opaque.⁷⁴

This defense of the thesis can be stated in the following manner. Begin with a robust, substantive conception of human nature. This conception can include facts about human nature that seem purely biological or descriptive in other ways. For example, the fact that we are born, that we have certain sorts of emotional and intellectual capacities that can be developed and exercised to varying degrees; that we age, and that we die. It also includes the fact we *need* certain things, which begins to sound less purely descriptive: that we need to be raised and cared for by other human beings, that we need shelter, food and drink, to exercise our bodies and minds, to learn a way of communicating with one another, and to live in communities with one another. And this conception also includes psychological observations such as the fact that human beings, or *normal* human beings, have an aversion to hunger, thirst, solitude, being deprived of bodily mobility, and death; or that human beings, or *normal* human beings, enjoy

⁷⁴ Other defenses rely on claims that are quasi-conceptual. As Scanlon puts the point in *What We Owe to Each Other*, a mere disposition to act, “does not in fact fit very well with what we ordinarily *mean* by desire” as “desiring something involves having a tendency to see something good or desirable about it” (p. 38). Others try to derive the guise of the good thesis from observations about linguistic patterns (by attempting, for example, to form Moore’s paradoxes for desires), or by simply relying on the analogy that desires aim at the good in the same way beliefs aim at the truth. These sorts of observations about what we ordinarily mean are helpful, but limited. When it comes to understanding the ways in which the thesis should be elaborated and why, it can be difficult to know how the investigation should go. Situating the thesis in this neo-Aristotelian background helps with this, as I will illustrate below.

pleasant bodily sensations, the use of their senses, and have some concern for other human beings. These latter claims are not wholly descriptive, nor are they biological; they are claims about what it is to be a *normal* human being, where the standard of normalcy here is not a statistical mean.⁷⁵ From this conception of human nature, or human life, we derive (in a way that I will not elaborate on here) a conception of the human *telos*: the human good. And once we have done this, we can compare particular human beings against this standard. Should a human being fail to achieve his telos, he is defective in a way central to his existence; should it seem like he is not aiming at his telos, he may not be a *human being*, at all, regardless of his biological species membership.⁷⁶

Invoking the category of *human nature* has a long and unhappy past, and it may be best to address a set of worries one might have right off the bat. For many, this language may bring to mind claims about the “unnaturalness” of masturbation, or homosexuality, or women working outside the home. Some might observe that it is a category most readily used by cultural conservatives, or those wedded to certain metaphysical views about our place in the universe that are no longer widely shared. Moreover, not only has this distinction been used in these dubious

⁷⁵ Though idea of “human nature” may have such connotations, Martha Nussbaum and Rosalind Hursthouse, emphasize that they do not mean to restrict their claims to the metaphysical.

⁷⁶ Consider, for example, a recent discussion of the guise of the good thesis by Matthew Boyle and Douglas Lavin in “Goodness and Desire”. They write:

On the [neo-]Aristotelian view, what underlies the fact that rational agents must desire under the guise of the good is a general point about the explanation of any self-movement or self-change... What these various kinds of explanation have in common is that they all have a teleological structure; and it is characteristic of the Aristotelian tradition to claim that, quite generally, this sort of structure applies only to a subject that is the bearer of a certain sort of *form*, a form that constitutes a standard of goodness for the subject in question, but that is equally implicated in any explanation of what the subject itself does.... This, it will emerge, is why rational self-movers [e.g. human beings] must desire under the guise of the good.

Boyle and Lavin conclude that the guise of the good thesis is best understood as true of the desires of good and non-defective people, but that “defective” desires are not under the guise of the good. Michael Stocker argues that the most plausible interpretation of the thesis is one against the background of a neo-Aristotelian framework, and concludes that the good, virtuous non-defective agent will desire under the guise of the good – but that the rest of us will not because we are defective. And Barbara Herman (1996), too, offers a general picture of how to think about the thesis when she incorporates Aristotelian elements into a Kantian framework to arrive at what she takes to be a more realistic conception of human psychology and development.

ways, it has also been used to justify atrocious forms of treatment of people who have simply been misunderstood. But though our understandings of human nature can be mistaken, though we are prone to set the limits of the intelligible at the limits of our imaginations, and though we often fail to understand people and cultures because of a failure to look more closely at what they are actually up to, these worries do not alone suggest that the concept does not, or should not inform our evaluative practices. We need to arrive at more *accurate* understandings of our natures, not to deny the sort of role such assessments play in our evaluative lives. Moreover, as I will discuss below, it is also a mistake to think that the unnaturalness of a person's desires or actions necessarily says anything about his *moral* character, or straightforwardly justifies any particular actions taken towards him at all. Similarly, a claim that something is natural – even if true – does not serve as a way of isolating it from moral objections, and the justified anger of others.

Returning to contemporary philosophical discourse, arguments that rely on constitutive norms of human nature are perhaps most familiar when it comes to questions about why human beings, regardless of what actual ends they may actually have, nonetheless always have reason to be *moral*. But one can see how the line of thought used in such arguments extends to investigations about the nature of our mental states, as well. In particular, it adds content to the idea that our parts, including our mental states, have certain aims, or goals – that they are meant to function in particular ways, independently of the goals that their particular bearers may have. Given our natures, and given what is good for us, our hearts are *meant* to pump blood through our circulatory system. Should a heart do this poorly, it is a *bad* heart; if it fails completely, it is no heart at all. Likewise, our beliefs are *meant* to get at truth, and our desires are *meant* to get at goodness. But unlike the failures of our organs, it is of central importance to our understanding

of ourselves as *rational agents*, who are responsive and intelligent in a way that non-human animal and plants are not, that our mental states like our beliefs and desires “get it right” through some kind of rational activity. This is the sense in which our desires are under a guise of goodness: if we end up desiring something that is in fact no good at all, our desires are making a kind of mistake central to their purpose; if a motivational state doesn’t seem responsive to considerations of goodness at all, it is not really a desire, but something else.

Understanding the guise of the good thesis in light of this neo-Aristotelian picture also explains what I called the “plausibility” of successful desirability conditions. Recall that it doesn’t help to learn that the mud-wanter wants the saucer under the guise of the intrinsic good of having saucers of mud, but that it *does* help to add a desirability condition like, “I want it because it’s medicinal”, or “I want to play with it”. This is because on this picture, the sort of teleological explanation that provides illumination here is one that references some substantive human good, or at least, something that a human being could plausibly see as a substantive good. It is not illuminating to learn that the mud-wanter wants the saucer of mud under the guise of the *intrinsic goodness of saucers of things* because we (currently) have no plausible conception of human life in which saucers of mud have that sort of status. As Talbot Brewer puts the point in his defense of the guise of the good thesis, “[The idea that actions are expressive of one’s conception of the good] carries the interesting implication that when we attempt to bring agency into view, we must assume a partial community of shared or at least mutually intelligible values”.⁷⁷

This basic picture is no doubt powerful. And indeed, I will accept much of it here. I will accept the indispensability of teleological explanations of human behaviour and normative conceptions of human nature, and I will accept that our mental states have irreducibly normative

⁷⁷ Brewer, *The Retrieval of Ethics*, p. 24.

and constitutive functions or aim. I will even accept that our desires, in particular, are sensitive to the desirability of their objects. My objection is to the thought that our desires have as their aim, the good. Rather, they are under the guise of the naturally attractive, and those things that are naturally attractive need not be, in any way, good. Should a human being desire something that is in fact *not* good or valuable, but is nonetheless naturally attractive, her desire need not be making any sort of mistake. There are two senses in which something can be “desirable”, and the things that it would be unintelligible to desire, the *adesirable*, are not identical to things that are undesirable.

V. THE DESIRABLE, THE UNDESIRABLE, AND THE ADESIRABLE

In order to decisively establish that conclusion, however, one would also have to decisively establish the truth of a complete conception of human nature, according to which some things that are naturally attractive to us (and so, the proper objects of well-functioning human desires) are not in any way good. Part of the difficulty of doing this is that the benefits of such a conception of human nature lie in its vast explanatory power, and its ability to make sense of aspects of our lives that are otherwise left mysterious. This project would involve a very broadly empirical outlook, and involve philosophical investigation paired with results from other disciplines, including other humanities. In short, it is no task for me to take on here.

So instead of proposing a complete conception of human nature, I will limit my discussion here to one potential aspect of human nature that is difficult to square with the idea that our desires aim for the good, rather than the naturally attractive, and which is very likely to be part of an accurate and complete conception of our nature. The aspect I will focus on here is that of human *aggression*. And in order to guide this discussion, it will be best to illustrate it with several examples. But importantly, these examples are not meant to serve as *counterexamples* to

the guise of the good thesis. Plenty of purported counterexamples to the thesis already exist, without having their intended effect: Stocker argues that a man might deliberately set himself on fire; a proponent of the thesis either references some good this man saw in doing this (the good of knowledge or, in certain contexts, the good of political dissent), or point out (rightly) that a person who did this “just because he wanted to” would be insane.⁷⁸

The point then, of such examples is not to add to this list. Rather, they will serve as a way of demonstrating how the guise of the naturally attractive thesis supplements our understanding of such cases in a way that the guise of the good thesis rules out. They illustrate that certain desires can be made fully intelligible with reference to a conception of human nature that includes aggression as major aspect of human motivation, and the evaluative qualities of the desired object – namely, that it is something naturally attractive to us. Moreover, given this conception of human nature, should a person desire something out of aggression, he may be doing so intelligibly and without suffering from a kind of illusion, *even if that thing is in fact not good at all*. But this does not mean that one can desire *anything* – there are substantive limits on what sorts of objects can be intelligibly desired, where those limits too, are informed by our conception of human nature. Desires and actions are under the guise of the *naturally attractive*, but not the good.

To see how these are in fact two different theses, I need to establish two claims: (1) that one’s desires can be rendered fully intelligible given some conception of human nature and the evaluative qualities of the object, and (2) that such a desire need not be mistaken (in a kind of quasi-perceptual way) when that object is not in fact good. Consider first an example from the television show, *Friday Night Lights*, which focuses on an American high school football team. In other words, it focuses on rather ordinary kinds of agents and rather ordinary desires. The

⁷⁸ Stocker (1979).

example we can use has, nicely, exactly the sort of shape that Anscombe's original case does.

Fullback Tim Riggins is asked about his reputation as one of the most aggressive players on the team, "What do you find appealing about that?" His answer is simply, "I just like to hurt people; I like to take them down". Now, given a plausible, substantive conception of human nature that recognizes *aggression* as an important aspect of human nature, Tim's desire is fully intelligible.

This conception of human nature can also serve to explain cases of desire and agency in which the agent is less aware of what has happened. Take for example, how Joel Feinberg, in a discussion on the unintelligibility of doing something bad for badness' sake, describes a personal experience in which he did something, but not under the guise of the good and so – for him – not something he can fully understand:

I had known Hal East [not his real name] for about a dozen years... Hal was remarkably kind and helpful to me and my family, arranging housing, transportation, introductions, and so on. For years after that, I saw Hal every couple of years, usually at social functions during the annual meeting of the American Philosophical Association...

One year I happened to meet Hal at one of the APA receptions... Suddenly I turned to Hal, and in all seriousness proceeded to insult him in a gross and vulgar way. ("Hal, you are a stupid son of a bitch.") Hal was taken aback and puzzled. After a considerable pause he asked, "Are you serious?" There was nothing at all jocular in my manner. That gave me a chance to recover, and I did. "No," I said. We both withdrew in a state of puzzlement, I am sure, and relief. But will either of us ever understand what had happened?⁷⁹

For Feinberg, this episode is just like the case of Anscombe's mud-wanter: he cannot understand his own motivations and behaviour in an intelligible way. He is not even able to attribute to himself the sort desire or intention that could make what he does seem like an action at all. This is because, as Feinberg suggests, there was nothing even *apparently* good about insulting his friend in this way.

⁷⁹ Feinberg, *Problems at the Roots of Law: Essays in Legal and Political Theory*, pp. 162-163.

But if this episode was simply an expression of aggression, there is a way to understand why Feinberg did what he did. On this interpretation, he wanted to hurt Hal's feelings, and he wanted to do that because human beings are naturally aggressive who find hurting one another's feelings attractive.⁸⁰ Thus, with this conception of human nature in mind, we are able to provide more possible, and plausible, understandings of our desires – rather than simply withdrawing in puzzlement from behaviour that can otherwise seem simply mechanistic, or compulsive.⁸¹ And it can serve to illuminate more general phenomena, as well: it is this picture of human psychology, for example, that motivates Nietzsche's suggestion that at the very heart of certain practices of justice is just the human intrinsic desire for domination and violence.⁸²

These observations are not yet enough to show that the guise of the good thesis is false. A defender of the guise of the good thesis may agree that human beings naturally desire violence and domination, and do so in ways that are fully intelligible. But when they do so, it is because such things *appear to be good* to human beings, though they are actually not. And so, when a person desires violence or domination, she is still desiring “under the guise of the good” – such

⁸⁰ I should emphasize that this is just *one*, possible interpretation of what happened here. There are even other explanations, even purely causal ones available.

⁸¹ This gives us an alternative reading of cases like this that is typically offered. In “For Badness’ Sake”, David Sussman treats Feinberg's actions as “perverse” – as done for the sake of badness. While I do think that such actions are possible as well (though I won't argue for this here), Feinberg may have bullied Hal for the sake of hurting his feelings. And this need not be done for the sake of badness, nor for the sake of some goodness there is in hurting other's feelings.

⁸² For example, his suggestion in *The Genealogy of Morality* that creditors are willing to accept their debtor's pain as *payment*, not as *punishment* for failing to pay back a debt, because it allows the creditor to experience “the pleasure of inflicting power on the powerless” and “the joy of sheer violence”. Nietzsche, *Genealogy of Morality*, p. 40.

things come to us as *disguised* as goodness, already.⁸³ To desire such things is to always already be subject to a kind of illusion.

No doubt this is one plausible interpretation of someone like Tim Riggins: on this interpretation, he is not just attracted to bringing about the pain of his opponents, but he also sees this as something that is *good*.⁸⁴ And it is plausible that a human being would think that violence is good, and so, have that appearance give rise to a desire to pursue it. I would not deny that this is a sound understanding of such cases, instead, I would insist that another is possible. While Tim *might* be suffering from an illusion, he might not be. Given the substantive conception of human nature proposed, his desire is sensitive to the sort of thing it is meant to be sensitive to. This is not just a *causal* explanation, in the way that referencing a burst of chemicals in his brain as the cause of his outburst. Rather, that tackling an opponent in a brutal way, or insulting one's friend in a "gross and vulgar way" would hurt those people is exactly the sort of thing that makes these actions naturally attractive; and so if one desires such things, one's desire is responding to exactly what it is that aggressive desires are supposed to respond to.

And so, someone like Tim Riggins may very well be aware that hurting people for the sake of doing so is not something *good* to do; he may very well know that it is something *bad*. Should we try to emphasize to him the various ways in which it is bad, and find that his desire does not diminish or show sensitivity to those considerations, is no evidence that his mental state is not a *desire*, in the way that should a person's belief-like state be utterly insensitive to

⁸³ A different response would be to insist that such things, given that we are naturally attracted to them, actually *are* good. But this seems, at least for some cases, a bad result. Moreover, even if those cases did involve some kind of goodness, it doesn't seem to be evidence for that claim that we are naturally attracted to such things.

⁸⁴ One might insist that the intelligibility of Tim Riggins' desire is due to the fact that *he believes* that hurting people is good, or that hurting people gives him pleasure and that pleasure is good. But this need not be the case, and in any case, would only salvage the version of the guise of the good thesis understood as a claim about an agent's evaluative beliefs – and as I've discussed above, there are other reasons for thinking that that particular interpretation of the view isn't right.

considerations that bear on its truth, this is some evidence that that state is not really a belief. Nor must we conclude that he is simply being irrational.

This is because his desire needs only to be sensitive to the naturally attractive, which, given a certain conception of human nature, causing pain in another is. If such a person is “defective”, he is defective according to moral standards, or prudential standards, or some other standard. Given this alternative conception of human nature on which our aggression itself aims at violence and domination, he is not defective as a *human being*, and his desire is not defective as a *desire*. Moreover, this conclusion would explain and reinforce the reaction that some of us may have to Tim’s answer: that it may be an honest expression of a person’s clear-eyed and perceptive understanding of his own (fully intelligible) desires, not simply the result of a stubborn illusion.⁸⁵

As mentioned, I cannot argue here that this conception of human nature, one that posits aggression as source of human motivation, is decisively correct. It is instead meant to illustrate how considerations of natural attraction diverge from considerations of plausible goodness a once we take it into account. But there are other ways to motivate this claim: one is that our reactions to one another seem to already implicitly assume this distinction. Consider, for example, one’s reactions upon learning that serial killer Jeffrey Dahmer not only murdered and sexually violated his victims, but also that he carefully labeled and stored their organs; or that as

⁸⁵ Nietzsche, apparently, finds amusing Augustine’s attempt to re-describe an event in which he and some fellow teenagers stole pears from a neighbor, seemingly for the *badness* of doing so. He writes to Overbeck:

...I was just reading the Confessions of Saint Augustine with great regret that you are not here with me. What a highfalutin wordsmith! Such tear-jerking phoniness! How hard I laughed, for example, over a “pear theft” of his youth, made the basis for his account of student days.

Augustine cannot understand how it could be in his nature to be attracted to this sort of mischief. After attempting many different explanations of why he did what he did, eventually decides that he stole the pears out of *the goodness of companionship*. Nietzsche, always quick to emphasize and even embrace the meaner parts of our nature, finds Augustine’s whole rigmarole amusing and “phony”, given what (according to Nietzsche) we are actually like. This passage is from Gary Wills, “Augustine and the Nature of Sin” (though he doesn’t cite where in Nietzsche’s letters it is from).

a child he was delighted by the sound of his father removing the bones of a dead animal from underneath their home. Rather than feeling inclined to resent or demand a justification from Dahmer for the things he did, one might find oneself withdrawing these attitudes and the emotional vulnerability they presume, and cultivating the objective stance toward him instead. There is something that we just cannot understand about Dahmer's mind and motivations, and this lack of transparency – so the defender of the guise of the good thesis may contend – is straightforwardly related to the *badness* of the things that he desired, and did. The *worse* the object of a person's desire, the less intelligible we will find him; the badness of what he seems to desire renders his motivations "beyond the pale".

But notice that one can have desires for objects that are much worse, in a number of important respects, than wanting to listen to the bones of a dead animal rattling together, but which are *more* intelligible as the objects of human desires. Again, if we accept that it is in human nature to enjoy hurting or dominating others, we can find fully intelligible – while morally objecting to – the desires of somebody who rapes and murders vulnerable members of society, or (much less seriously) who desires to bully or belittle his friends or colleagues. And we understand these people more than we understand a person who desires to hear the sound of rattling animal bones, or endlessly count blades of grass, or drink cans of paint, even though those particular motivations – though unquestionably *weird* – are not for things that are *worse*.

We have so far seen that there are plausible conceptions of human nature that support the thesis that we desire under the guise of the naturally attractive, and that this is distinct from the claim that we desire under the guise of the good. Furthermore, the dimensions of our reactive attitudes, and in particular, when it is fitting to take on the objective stance towards someone, better cohere with this result. And there is a third piece of evidence in favour of marking the

distinctions as I have done so here. In Book VII of the *Nicomachean Ethics*, Aristotle observes that along with *akrasia* and *vice*, another bad sort of state for a human being to be in is *thêriotês*, translated as the “bestial”, “beastly”, or “brutish”.⁸⁶ Such states, he suggests intriguingly, are “less grave than vice, but more *frightening*”.⁸⁷ Another notable feature about this category of evaluation, as Pavlos Kontos notes, is that Aristotle’s examples of beastly desires are “puzzling”, in that – even keeping in mind cultural differences between the ancient Greeks and ourselves – they “hardly have any moral relevance”⁸⁸. Indeed, they do not seem to exist along one dimension of *badness*, moral or otherwise: Aristotle includes examples of things that are terrible, such as ripping a child from a womb; but he also notes relatively benign things like plucking hairs and eating dirt.

We can make sense of Aristotle’s categories in the follow manner. When a person’s desires are beastly, or when she performs a beastly action, we cannot understand what she is doing, and because of this, we cannot evaluate or react to her in the same ways that we can evaluate a person who desires or does something that we find intelligible to desire or do, even when that thing is bad, and her actions vicious. Though the distinction between the human and the bestial is *important*, it should not be conflated with a distinction between the moral and the immoral, or the virtuous and the vicious. While it may be crucial for our evaluations of a person that we find her motivations intelligible, it may not at all reflect a person’s bad *moral* character

⁸⁶ None of these terms, I think, are ideal: all suggest that we should link this state with non-human animals. But the distinctions, at least as I am conceiving of them, can be applied to some non-human animals – a cat can intelligibly desire to chase a mouse or bat at a shadow, but cannot intelligibly desire to eat shards of glass or its own tail. Perhaps our ability to apply these categories depends on how similar their natures are to our own: it is not as clear how one should apply them to very simple reptiles, or insects.

⁸⁷ NE 1150a37, my emphasis.

⁸⁸ Kontos, “Non-Virtuous Intellectual States in Aristotle’s Ethics”, in *Oxford Studies In Ancient Philosophy*, XLVII, Brad Inwood (ed.), p. 224, fn. 39.

that she has unintelligible ones. As Aristotle writes, “One sort of vice is human, and this is called simple vice; another sort is called vice with an added condition, and this is said to be bestial or diseased vice” (1149a17). Human or “simple” vice, though indeed sometimes directed towards things that are bad, is less opaque to us than bestial, or diseased “vice”.⁸⁹ And we react to a beastly person in a way that reflects this: we will not characterize or assess her behaviour in a morally robust way; she is not subject to the same attitudes, such as indignation or resentment, that are natural responses to (humanly) vicious desire or action. Aristotle observes that we react to the beastly person with *fear*; Strawson would add other, non-reactive emotions such as revulsion, and pity.

A defender of the guise of the good thesis may respond that she is able to account for this category as well, perhaps by suggesting that the bestial just includes those things that are not apparently good, at all. Though I have no exegetical motivations in this paper, I nonetheless think there another interpretation is available. Consider again the list of items that Aristotle uses to illustrate the category of the bestial, where what is salient in each of the examples does not simply seem to be a shared lack of value. The beastly person has desire-like attitudes towards the adesirable, but the adesirable is not simply all those things that are not plausibly *good*. They are instead things that Aristotle deems not “naturally pleasant” to human beings, or what I have been

⁸⁹ There has been relatively little written on this aspect of Aristotle’s moral psychology. But in “Aristotle on Brutishness,” John Thorpe has reached a conclusion that is very similar, I think, to the one that I have suggested here. On his reading of Aristotle, “...normal people have normal desires; but desires can diverge from the norm in two ways, either in their intensity, or in the *naturalness of their objects*. Sinners with normal desires function in the ethical realm of vice; those with unusually intense desires function in the realm of akrasia; those who desire abnormal objects function in the ethical realm of brutish vice; and finally, sinners whose desires are abnormal in both ways function in the realm of brutish akrasia” (p. 682). Thorpe makes a point to emphasize that “normal” here neither means statistically normal, nor non-vicious, as it is precisely the “sinner” who has normal, but vicious desires. He also suggests Aristotle means to provide us with what he calls a “polymetric” system of evaluation.

calling “naturally attractive” to human beings.⁹⁰ The bestial person is not “beyond the pale” because what he wants is extremely, horribly *bad* – it is just not naturally attractive. The person with a vicious desire, in contrast with a brutish impulse, wants something that is naturally attractive, although it may be something that is naturally attractive and *bad* – indeed, it may be much worse than what the brutish person wants. While Aristotle, given certain other of his commitments, has difficulties accounting for how this sort of vicious person is able to act in a clear-eyed manner⁹¹, with a particular conception of human nature in mind, we are able to fill in the details.

Finally, although I have discussed a guise of goodness as a necessary condition on the intelligibility of human desire and action, I now want to consider an implication of the position I have argued for here. Though I will not fully defend this claim, it can help illuminate further how it is that considerations of natural attraction come apart from considerations of goodness. The implication is that not only is a guise of goodness not necessary for intelligibility, in some cases it isn’t *sufficient*. This is easier to notice in cases where we are suspicious that a particular person could really desire or be doing something good, for *that* reason, given what we know about her. If she is, for example, a particularly conniving or tyrannical sort of person, we will not find plausible the idea that she is truly motivated by the guise of somebody else’s good. If we’re to find her intelligible at all, we may have to interpret her as acting in a way consistent with the kinds of desires we know she has – “she’s buttering you up,” “she’s manipulating you”.

But if what I have argued here is right – that given a certain understanding of human nature, the distinction between the naturally attractive and unattractive is orthogonal to the

⁹⁰ I am not endorsing the claim that the naturally attractive is also literally “pleasant”. There are forms of pain that are naturally attractive to human beings – such as the pain that accompanies rigorous physical activities or achievements.

⁹¹ In particular, of his suggestion that agents who act against their best judgment have only a degraded form of knowledge or grasp of the relevant evaluative facts. NE 1147a10–24.

distinction between the good and the bad – this point may be generalizable. There may be some objects that we find plausibly *good*, but not intelligible for human beings to desire because not naturally attractive. If we came across a person who reports having a desire for a plausibly good, but undesirable, thing, we would experience the same sort of withdrawal of the reactive attitudes – but this time, the *positive* ones. One example may be of a person who sincerely “desires” to determine through rigorous and methodologically sound empirical testing whether pigeons prefer to eat red bell peppers or green bell peppers. Another may be of a person who “desires” that the world’s mosquito population not go hungry with a successful global distribution of mosquito netting. In such cases, a person may be able to easily cite what strikes them as good about what they desire, and we may even find such citations plausible: for example, it’s for the good of scientific knowledge, or for the good of mosquitos. But nonetheless we may get the sense that we don’t understand what it could mean to truly *desire* these things, just as we don’t understand what it means to truly desire a twig of mountain ash.

In more realistic cases, we feel an urge to learn more about *why* somebody wants what they say they want, even when we think that what they want is good. When we hear, for example, of a couple’s desire to adopt as many children as possible, or a person’s desire to take a job as a high-paid corporate lawyer in order to give more money to charity, we may be able to quite clearly see what is *good* about what these people want to do.⁹² But rather than straightforwardly praising them, one might experience a kind of befuddlement, or estrangement, instead. We are confused; we need to hear more. This is not to suggest that such people are not *worthy* of praise, or that this confusion cannot be alleviated. It is instead to suggest that we don’t

⁹² Larissa MacFarquhar’s recent book, *Strangers Drowning*, provides in-depth profiles of what she calls “extreme do-gooders”. That we find these people puzzling is a premise of the book; but we need not find them puzzling because we do not believe that what they want to do is *good*. Rather, at least one explanation of why we find them puzzling because *given human nature*, it is hard to understand why they want what they want.

know yet what to think about them: it is to acknowledge a tension that should simply not arise should we think that the good just is naturally attractive.

VI. CONCLUSION

There is one important motivation that drives at least some people to deny the guise of the good thesis, which I have not yet explicitly discussed here. It is the thought that not only is it in the nature of our desires and our emotions to resist rationalization, but that *we* should resist rationalizing them. One might find appealing – not simply for meta-ethical reasons, but for quasi-ethical ones as well – the thought that sometimes a person simply wants what he wants and that to accuse him of making a *mistake* in wanting what he wants is not only to misunderstand the nature of desire and its relationship to reason, but to try to level a particular kind of criticism of him that is misguided, or even excessively moralistic.⁹³

But we should not conflate these two different interpretations of the relationship between our desires and reason. Doing so encourages two camps: one that treats desires as insensitive to evaluative considerations, and one that argues that we *must* treat desires as sensitive to evaluative considerations, but which then assumes that the desirable is synonymous with the good. And here again, a person with the sort of concern just mentioned might balk: sometimes I simply *want what I want*, not what is *good*. But, the rejoinder goes: you cannot “simply want” saucers of mud, or twigs of mountain ash; you cannot “simply want” to drink cans of paint, endlessly count blades of grass, or rub your face along the northeast sides of buildings.

⁹³ This seems to underlie, for example, Harry Frankfurt’s various discussions on love in *Contours of Agency: Essays on Themes from Harry Frankfurt*. For Frankfurt, it is not “unreasonable” to love (and desire, etc.) things that are (morally) bad because they are bad – if it is unreasonable to do so, it is because of the consequences that doing so will have on other things that one cares about.

I have argued here that there is a third way available to us that can accommodate concerns raised by both sides. This is to accept, *contra* the Humean, that desires have desirability conditions: there are limits to what sorts of things human beings can desire, because there are limits to what sorts of things human beings can be attracted to, and those limits are sensitive to the evaluative qualities of those things. A person who “desires” something that is undesirable isn’t exactly *irrational*, but she is opaque to us in a very important way, nonetheless – and we must be able to make sense of this opacity. And the limits on what sorts of things are desirable are not simply accidental – it is not as though to determine such limits we need simply to observe what human beings exhibit motivated behaviour towards. These constraints are, I’ve argued, informed by a substantive conception of human nature.

However, this position denies, *contra* defenders of the guise of the good thesis, that everything that is desirable for human beings is in some way even plausibly good, or that all human desires aim at the plausibly good. A person who wants something that *isn’t* plausibly good may be imprudent or cruel, but as long as it is something that is naturally attractive to human beings, he might not be making a mistake in desiring it, nor failing to desire at all.

But I have also hoped to emphasize that there is much to be gained from taking seriously the concerns of defenders of the guise of the good thesis, in particular, from its neo-Aristotelian adherents. They are, as I have argued, right to emphasize the normative limits of human desire. They are also right to suggest that when we consider what things are good, and how it is best for us to live, it is important that we take into consideration a substantive, realistic account of human nature and psychology. However, though both our conceptions of human nature and our conception of the good will no doubt be informed by one another in certain important respects, we should not assume that harmony between these two conceptions is somehow guaranteed – as

though all that is needed is some improvement in our empirical circumstances to forge a link between us, and goodness. Rather, we must take into consideration the real possibility that, given our natures, the good is not simply made for us, nor we for the good.

Boredom & The Disharmonious Mind⁹⁴

I. INTRODUCTION

I've argued that our desires, when intelligent responses to the world, need not be intelligent responses to considerations about goodness. I will now bolster the two main conclusions that I have been arguing for: that akrasia is psychologically coherent, and possible, and that it is not always a bad thing that a person is akratic. I will engage directly with the Aristotelian thought that psychic harmony is necessary for virtue, and I will focus on forms of akrasia that have been under-explored: those that are best explained by a person's experiences of boredom and interest. I will conclude by suggesting that while wholeheartedness may be characteristic of certain virtues, we need a more pluralistic approach to the virtues in general.

II. BOREDOM & PSYCHIC DISHARMONY

According to a standard conception of virtue, rooted in Aristotle, the virtuous agent not only knows what action would be all things considered best to perform⁹⁵, but in lacking "base

⁹⁴ This chapter previously appeared as an article in *Res Philosophica*. The original citation is as follows: Vida Yao, "Boredom & The Divided Mind," *Res Philosophica* 4 (2015) 937-957 (Special issue on Virtues and the Emotions).

⁹⁵ Hereafter, simply whatever action would be "best" or "right". Importantly, I do not mean to mask the fact that there important reasons (even within the domain of morality, which I will not be limiting my discussion here) for not conflating the language of good and bad with the language of right and wrong. See, for example, Robert Adams' discussion in *A Theory of Virtue: Excellence in Being For the Good*, and Susan Wolf's "Above and Below the Line of Duty".

appetites” and finding “nothing pleasant against reason”, he performs the best action *wholeheartedly*.⁹⁶ This unity of his motivations and emotions is taken to be the feature that distinguishes him from the enkratic or continent agent, who performs the best action but not without overcoming some reluctance in doing so. More generally, there has been a long philosophical tradition of viewing inner harmony as a mark of the ideal human being; in contemporary philosophical literature, one way in which this tradition has taken shape is by a tendency to regard inner harmony or coherence as essential to rationality, or by taking the standards of virtue and the norms of practical reason to coincide with, or reinforce one another.⁹⁷ Against this background, genuine excellences of character, so it seems, cannot be constituted by or based upon an agent’s motivational or emotional divergence from what she knows would be best to do: the gaps that can arise between one’s awareness of that action and one’s motivations to do it is precisely where vice or weakness take hold of one’s character.

In spite of its attractions, I shall argue that we need to reconsider whether wholeheartedness deserves this lofty status. I suspect that our theoretical attraction to it has been encouraged by first of all, an overly narrow focus on what it takes to be a paradigmatically excellent *agent* when considering what it takes to be a paradigmatically excellent *person*⁹⁸, and second of all, by taking for granted that cases of motivational or emotional ambivalence toward

⁹⁶ NE1152a.

⁹⁷ See, for example, Michael Smith’s *The Moral Problem*. In “Does Rationality Consist in Responding Correctly to Reasons”, John Broome argues for a more restricted claim that he calls the “enkratic condition”: rationality requires that an agent intend to do what she believes she ought to do. And R. Jay Wallace, in “Three Conceptions of Rational Agency”, puts the point this way: “If agent A has reason *r* to perform action X, and A is properly aware that *r* obtains, then A must be motivated to do *x*, on pain of irrationality” (pp. 217-18).

⁹⁸ In other words, our conception of a good person has been too heavily shaped by one’s ability to successfully perform the right action in the right way. This might be the result of moral philosophy’s predominant interest in providing theories that are “action-guiding” – an ambition that I believe has been made too central to ethical theorizing.

what one knows would be best to do are always indications of one's irrationality. Given these standard assumptions of philosophical investigations in practical reason and moral psychology, we too hastily assume that there is nothing left worth either exploring or appreciating about the distinct ways in which a person's psychology can be divided.⁹⁹

I propose we look again. When we examine more carefully why it is that some souls are divided, *in the particular way* in which they are divided, we will find that ambivalence toward the best action need not be evidence of an undesirable or unattractive quality in a person's character. On the contrary, as I have argued in chapter 1 and as I will reinforce here, this disharmony may be the very foundation of, or partly constitutive of, certain qualities of character that we enjoy and appreciate in one another, and which we might ourselves wish to possess.

Here, I shall focus on another explanation for this disharmony, which involves the contrasting emotions of boredom and interest. Both emotions are notable for the ways in which they tend to influence our motivations, and – important for my discussion here – for the ways in which they can lead us to act, or be inclined to act, against our better judgment. In particular, we can be reluctant to perform those actions we know would be best to perform precisely because of how boring we find them; and out of occurrent feelings of boredom, we can become interested in or attracted to performing actions we don't think would be good to perform at all. But I shall argue that given the relationship between one's patterns of boredom and interest, and certain positive character traits, the bored enkratic or akratic may be no less excellent given this disharmony in her soul. While there are many negative character traits commonly attributed to

⁹⁹ One important difficulty in stating the difference between the disposition of virtue and the disposition of continence is that given a plurality of values, we would expect the virtuous agent – in being sensitive to this plurality – to be divided among conflicting but good options. My arguments here will not focus on emphasizing how the virtuous agent might be internally conflicted given a plurality of value. So a more precise way to characterize the difference between the virtuous agent and the continent agent would be to emphasize that the continent agent is ambivalent toward performing the best action, in spite of there being no, or comparably little, competing *value* in the other option available to her.

those that are easily bored – such as shallowness, or a lack of gratitude or appreciation – there are a number of positive ones that are perhaps best exhibited by bored enkratics or akratics, in particular.

In light of the fact that there seem to be other genuine virtues that necessarily depend on wholeheartedness even in the face of extreme boredom, these observations should not lead us to the conclusion that being divided is somehow *better* than being wholehearted. But should the conclusions I draw here be compelling, neither should we accept the converse. Instead, I hope my discussion encourages moving away from an understanding of the virtues that lends itself to this sort of ranking, and toward adopting a more flexible, or pluralistic approach to delineating and describing the virtues: one that allows for the appreciation and celebration of virtues that may conflict with the requisites for others.

III. WHOLEHEARTED ACTIVITY

Let's begin by taking into consideration what this particular kind of wholeheartedness or unity looks like when instantiated in a person, and some reasons for why the standard conception of virtue emphasizes its importance. We can start with a suggestion from Julia Annas, who writes that the distinction between the disposition of virtue and the disposition of continence, self-control, or “*encrasia*”, is one that we should be familiar with. According to Annas, it is:

...the everyday contrast between someone who does the right thing, but has to battle with his feelings to do so, and thus acts reluctantly and with a sense of pain and loss, and the person who does the right thing and whose feelings endorse the action, and who thus acts gladly and with pleasure. We all do recognize (mostly in our own case) the difference between the merely self-controlled... and the person who does not have to be self-controlled. And we take pleasure or the absence of it in acting to be the chief mark of this distinction. Aristotle draws the distinction in

these terms, but it is clearly present in later authors even if they do not use the word, since it is regarded as a matter of common sense.¹⁰⁰

Granting that this is a familiar and commonsense contrast, we can nonetheless imagine at least two different interpretations of what it would take to maintain unity, or wholeheartedness, in doing the right thing. The first would be to take Annas at her literal word here and conclude that the virtuous agent always experiences pleasure, or positive interest, in doing the right or best thing. The enkratic agent, in contrast, experiences no pleasure in doing the right thing – and this is why she must exercise self-control in order to do it.

The worry with this first interpretation is that there are cases in which we might think that the virtuous agent shouldn't take *positive pleasure* or *interest* in doing the right thing. Taking Old Yeller out back to be shot might be the right thing to do, but it would seem callous or insensitive, if not insane, should one literally “act gladly”, or take pleasure in doing this. So we should supplement this initial interpretation in order to accommodate cases like this, while still preserving the thought that is a more “wholehearted” way of shooting Old Yeller, and a less wholehearted way.¹⁰¹ I propose the following: depending on the details of the particular action in question, while the enkratic experiences *pain*, or *reluctance*, or some other negative reaction that must be overcome, the virtuous agent – though not experiencing *positive* interest or pleasure or enthusiasm – at least does not experience these negative reactions, or at least does not experience them to the extent that she must rely on self-control in order to do what is right.

Now, why think that wholeheartedness is a defining mark of virtue? One rather straightforward reason is that the wholehearted agent will more reliably perform (or at least

¹⁰⁰ Julia Annas, *The Morality of Happiness*, p. 53.

¹⁰¹ One might think that the distinction between the virtuous agent and the enkratic agent collapses in cases like this – but if this is so, we'll already have some reason to think that wholeheartedness is just not as essential to virtue as the standard conception maintains.

attempt to perform) the best actions, for the right reasons: she is just less susceptible to acting against her better judgment, while the enkratic seems to always skirt the edge. But another consideration is that the wholehearted agent will not be subject to the inner turmoil and pain experienced by those whose emotions are less unified with their knowledge of what actions are best to perform, and the avoidance of this tension makes the disposition of wholeheartedness preferable. As Annas puts the point, it is this tension that renders the state of self-control, or *enkrasia*, a “lower stage” than the state of virtue, even though both the virtuous and the self-controlled agent may ultimately perform the very same actions. She writes, in approval:

What the ancients stress [in stressing the importance of wholeheartedness] is just the common thought that conflict and stress are signs of something’s failing or going wrong, and that a state where these are absent is preferable to a state where they are present. Virtue is not just different from self-control; the harmony in the virtuous between action and feeling makes it *preferable* to self-control.¹⁰²

In contrast to the inner harmony of the virtuous agent, the enkratic is susceptible to at least three distinct forms of psychological disharmony: first, she is motivationally divided between the better and the worse in the first place; second, she must exercise self-control in order to do the thing that is better; and third, because she is motivated to do the worse thing to begin with, she is likely to feel some regret or dissatisfaction even when she successfully does the better thing – the discontent that one can feel in turning down, for example, a desired drink or helping of dessert, even while knowing that it was better that one did so.

The virtuous agent, having no motivation to do the worse thing in the first place, gracefully transcends this inner conflict and stress. As John McDowell describes her, she

¹⁰² Annas, *The Morality of Happiness*, pp. 53-54.

renounces other options “without struggle”, achieving a kind of alluring “sublimity” or “serenity”.¹⁰³ Furthermore, we may even think of her as possessing substantive qualities of character, such as autonomy, self-confidence, sincerity, integrity and a distinct lack of self-alienation, in virtue of the harmony that resides within her soul.¹⁰⁴

Moreover, there is, as mentioned, a widely shared assumption that there is an important connection between rationality and inner harmony. Rosalind Hursthouse, for example, emphasizes the desirability of this inner harmony by making explicit its connection to norms of practical rationality, and the connection between practical rationality and human nature. Denying that the virtuous agent’s excellence over the merely continent agent consists in the reliability of his performance of the best action, she writes:

The Aristotelian view of human nature is that, *qua* rational, it can be perfected by getting our inclinations into harmony with our reason. If my inclinations are not in harmony with my reason, and if getting them into harmony is something that human rationality can achieve, then the people whose inclinations are in harmony are, *ceteris paribus*, better human beings, closer to excellence (virtue), than I am.¹⁰⁵

Nonetheless, as I proposed earlier, we need to look more closely at the distinct ways in which a person can fail to be wholehearted before concluding with Annas, Hursthouse, McDowell and the ancients, that inner harmony is not only different from but in itself preferable to or more excellent than inner division.¹⁰⁶ Echoing Aristotle, while there is only one way for a

¹⁰³ John McDowell, "Are Moral Requirements Hypothetical Imperatives?" (p. 27-28).

¹⁰⁴ One need not maintain a rationalist picture of human psychology to take this inner harmony to be essential to a good human life. See, for example, Harry Frankfurt’s discussions on the importance of *wholeheartedness* in *The Reasons of Love*.

¹⁰⁵ Rosalind Hursthouse, "Are Virtues the Proper Starting Point for Morality?" (p. 104).

¹⁰⁶ It might be suggested that we admire the person who must struggle to do the right thing *more* than the person who need not struggle, given that it shows an admirable strength of will or a full sensitivity to the considerations that weigh against him. This is a different line of thought than I will pursue here. Annas rejects this proposal, suggesting that although we do admire the disabled for overcoming their disability,

person to be virtuous there are a variety of ways in which human beings can miss the mark, and our evaluations of a person will differ depending on the particular way in which he falls short of this standard in particular.¹⁰⁷

In philosophical discussions of this variety, the opposing experiences of *boredom* and *interest* have been largely overlooked as explanations of why an agent may not be wholeheartedly motivated to perform what she knows to be the best action. Typically, our attention focuses on the misalignment or disproportionality of one's appetites, or on "hotter" emotions like anger and fear. This should strike us as a surprising omission once we reflect on how, in our everyday lives, boredom typically effects our motivations.¹⁰⁸ It seems that we are often disinclined to do what we know would be best simply because it is also something boring to do: we halfheartedly wait at the DMV to renew our licenses, politely endure conversations with the self-absorbed, and plod through the routine steps it takes to maintain an orderly home or office. And similarly, it seems that we are often positively *interested* in doing things that we

it would be confused to infer from this that the state of being disabled is preferable to, or as preferable as, the state of being able-bodied.

¹⁰⁷ "... there is only way to be correct. This is why error is easy and correctness is difficulty, since it is easy to miss the target and difficult to hit it. And so for this reason also excess and deficiency are proper to vice, the mean to virtue; 'for we are noble in only on way, but bad in all sorts of ways', " (NE1106b28-33). Aristotle was sensitive to the fact that our evaluations of one another will differ depending on the way in which we fail to be virtuous: he proposes, for example, that being overcome by one's appetites is more shameful than being overcome by one's spirit (NE 1149b).

¹⁰⁸ What is boredom? I won't attempt to provide an analysis of it here. Instead, we can borrow from Wendell O'Brien's, recent article, "Boredom". He writes that boredom is:

- (1) a mental state of
- (2) weariness,
- (3) restlessness, and
- (4) lack of interest in something to which one is subjected,
- (5) which is unpleasant or undesirable,
- (6) in which the weariness and restlessness are causally related to the lack of interest.

My discussion will emphasize the consequences of features 3-6.

simultaneously recognize are of very little value, or no value at all, to do. Think of a person who knows that it would be best to get a good night's rest before an important meeting she has in the morning, but finds herself engrossed by the inane television show she just put on. She may not believe that the show itself is any good, or that it would be good for her to continue watching it, but nonetheless be absorbed enough to stay up an hour or two longer than she knows would be best. Furthermore, one notable feature of the experience of boredom is its tendency to give rise to motivations to do things that we simultaneously recognize are not good to do, at all. Boredom renders us restless; out of boredom, people eat though they are not hungry, fiddle with things until they break, vandalize just for the sake of it, and pick fights with one another.

But perhaps this omission is justifiable. Perhaps these cases are not really best understood in the way that I've described them – as instances in which considerations of interest and boredom conflict with considerations of value and disvalue, respectively. This thought would be consistent with the particular kind of boredom that tends to draw both philosophical and literary attention: one that is typically associated with the experience of *disillusionment* or *disenchantment* that accompanies a sort of widespread evaluative nihilism.

For example, this is how we might interpret Henry, the narrator of John Berryman's poem, "Dream Song 14", who begins by stating that, "Life, friends, is boring", before elaborating:

...Peoples bore me,
literature bores me, especially great literature,
Henry bores me, with his plights & gripes
as bad as Achilles,
who loves people and valiant art, which bores me.
And the tranquil hills, & gin, look like a drag
and somehow a dog
has taken itself & its tail considerably away
into the mountains or sea or sky, leaving

behind: me, wag.¹⁰⁹

Henry's boredom, we might think, is a result from his seeing the world as empty of anything of genuine value. Even supposedly "great" literature turns out not to be so great, after all. And his general withdrawal from the world of value, we might think, just *is*, or is essentially characterized by, the experience of boredom.

If this were the only sense of boredom familiar to us, then it would be natural to conclude that an agent cannot be properly understood as being bored by what she really believes is best, or is best, for her to do: that it cannot sustain her interest is itself a sign that it *lacks* value for her. Harry Frankfurt suggests something like this in his discussion of why it is particularly *bad* for human beings to be bored. When we are bored, Frankfurt claims, this is an indication that we do not find the objects of our boredom *valuable* or *important*.¹¹⁰ This suggestion nicely pairs with and is explained by a more general position that what *gives* something value is a person's finding positive interest in it.¹¹¹

Consider an implication of this understanding of the relationship between boredom, interest, and value. It suggests that a person's life would be unquestionably improved if she were,

¹⁰⁹ Berryman, John. "77 Dream Songs."

¹¹⁰ As Frankfurt writes, "We do not care about any of it; none of it is *important* to us. As a natural consequence of this, our motivation to stay focused weakens; and we undergo a corresponding attenuation of psychic vitality", *The Reasons of Love*, p. 54.

¹¹¹ "It is true that the beloved invariably is, indeed, valuable to the lover. However, perceiving that value is not at all an indispensable formative or grounding condition of the love. It need not be a perception of value in what he loves that moves the lover to love it. The truly essential relationship between love and the value of the beloved goes in the opposite direction. It is not necessarily as a result of recognizing their value and of being captivated by it that we love things. *Rather, what we love necessarily acquires value for us because we love it*," *The Reasons of Love*, p. 38.

no matter her actual circumstances, *unborable*.¹¹² If her positive interests and cares can themselves give rise to and create value for her, then a sustained interest in *anything* would create and maintain one's contact with some form of goodness.

But rather than being clearly attractive, there is something unsettling about the idea of being unqualifiedly unborable that this picture seems to leave unexplained. Consider John Rawls' example of a mathematician who takes pleasure in counting blades of grass in "various geometrically shaped areas such as park squares and well-trimmed lawns".¹¹³ Rawls points out, quite rightly, that we would find such a person, as he puts it, "surprising". And he is surprising in a particular sort of way: it is difficult to imagine what it is like to *be* him, persistently engaged in this activity. There are a number of features about grass counting that we might cite in an attempt to explain our puzzlement. Undoubtedly, part of the explanation may be that this activity is pointless, or valueless. But the strangeness of the grass-counter cannot be wholly explained by these considerations alone. Stamp collecting and jigsaw puzzle solving may be pointless, but we understand a person's interest in doing either, even in light of his and our awareness of the pointlessness of these activities. And there may be nothing valuable about smashing the icicles that have been growing on the windowsill, or pulling a mean-spirited prank on one's colleague, but again, it is no stretch of imagination to understand a person being positively engrossed by these activities, in spite of their lack of value. Counting blades of grass stands out as distinct and

¹¹² I borrow this neologism from David Foster Wallace's novel, *The Pale King*, in which a character remarks: "To be, in a word, unborable.... It is the key to modern life. If you are immune to boredom, there is literally nothing you cannot accomplish". Wallace presents the view that this is not only the key to modern life, but constitutes a kind of virtue: "Enduring tedium over time in a confined space is what real courage is". I'll elaborate below on the particular sort of admirable "unborability" that Wallace seems to have in mind and why it differs from the form of unborability that I will discuss first, and which is more normatively suspicious.

¹¹³ *A Theory of Justice*, p. 42.

surprising because not only is it in itself pointless and valueless to do, it is also – if any activity is – boring to do.

Given the dullness of grass counting, we might wonder whether Rawls' mathematician could really be fully attending to what he is doing, or whether his concentration is actually directed toward something else. Maybe his mind is elsewhere: contemplating proofs, or recollecting childhood memories. And so, although some small part of his mental energy is focused on the actual task of grass counting, the bulk of his attention is really directed elsewhere. Or perhaps he is – as Rawls suggests – a neurotic who uses grass counting to avoid the frustrations of engaging with other people.

But if we cannot explain his ability to sustain an interest in counting blades of grass in these ways (which each suggest that it isn't the task *itself* that he actually finds engaging), we may begin to suspect that – unless there is something interesting about grass counting that we haven't taken into consideration – it is his *experience* of that activity that is impoverished. This is because, as Bernard Williams has put the point, the experience of boredom can be “not just a tiresome effect, but a reaction almost perceptual in character...”¹¹⁴ And given this, it can – like other emotions such as anger and fear – have what amount to veridicality conditions: that is, a person's experience of boredom can be an indication that she is accurately perceiving the features of her situation or her activity, and furthermore, that in some cases, her *failing* to find certain things boring can suggest a lack of perceptual connection with her environment. In these latter cases, either one's attention must really be directed elsewhere (again, perhaps turned inwards), or one is failing to experience the activity as it really is.¹¹⁵

¹¹⁴ Williams, "The Makropulos Case: Reflections on the Tedium of Immortality," p. 95.

¹¹⁵ By claiming that boredom is perceptual, or “quasi-perceptual”, I mean to suggest that – at least for well-functioning human beings – there it is sensitive to the features of one's circumstances and because of

Williams' own discussion of boredom incorporates this quasi-perceptual aspect of boredom by emphasizing how the experience of boredom can kill off a person's desires to engage in the world by being an experience that is *responsive* to the features of her circumstances. On his reading of the case of Elina Makropulos – a woman who is given the capacity for an immortal life – EM doesn't become less and less invested in her life for no reason, as if her interests and cares have simply petered out over time. Nor is her boredom necessarily a matter of her no longer finding things to be of *value*. Rather, she becomes more and more emotionally withdrawn because after living too long as herself she senses that, "in the end, *it is the same*".¹¹⁶ And it is this thought, Williams thinks, that not only explains but also *justifies* EM's boredom with her unending life.

I will not evaluate Williams' argument that it would be intolerable because too tedious for *any* human being to be able to maintain her identity while living an unending life. Instead, I want to highlight just two features of his discussion that are germane to mine. First, that boredom isn't *blind*: it is an experience that one can have in response to the features of one's situation. This explains why there is, at least without further specification, something unnerving about the idea

this, there are *limits* to what it is that we can find interesting. This differs from a standard dispositional analysis of an emotion, because it involves a particular kind of normative standard – a person's experience of boredom can be *merited* or not, given features of her circumstances or her activity. While I am committed to any particular theory of the emotions, there are some that seem consistent with what I suggest about boredom here. For example, in "The Significance of Recalcitrant Emotion (or anti-quasijudgmentalism)", Justin D'Arms and Daniel Jacobson offer the view that "base emotions", such as amusement, anger, contempt, and disgust as the products of "relatively discrete special-purpose mechanisms that are sensitive to some important aspect of human life" (p. 138). Jealousy, for example, "monitors the social environment for potential losses of affection or allegiance" (p. 140). On this picture, we can understand interest and boredom as monitoring one's environment for things like novelty, and repetitiveness. For D'Arms and Jacobson, this picture of the emotions supports the thought that such responses are governed by what they call "norms of fittingness": "considerations of fittingness are all and only those considerations about whether to feel shame, amusement, fear, and so forth that bear on whether the emotion's evaluation of the circumstances gets it right: whether the situation really is shameful, funny, fearsome, and so forth" (p. 132). Something like this norm is what I will rely on when I discuss the idea that a person's boredom, or interest can be "veridical", or "justified".

¹¹⁶ "The Makropulos Case: Reflections on the Tedium of Immortality," p. 82 (my emphasis).

of being *unborable* – constantly engaged and fascinated, *regardless* of the features of one's circumstances. We seem to at least want some assurance that if one is unborable, it is because one is engaged with things that actually merit one's interest. Second, that boredom need not always be a response that one has in recognition of a lack of *value*. Rather, it can simply be a response that one has when what one is doing is no longer *interesting* – perhaps it has become lifeless and insipid in the way that an excellent song can fail to move us simply because we have heard it too many times.

Importantly, this allows for the possibility of a person's finding boring the very thing that she simultaneously recognizes would be good, or even best for her to do. Her present sense that a certain activity is boring need be no indication that it isn't something worth doing, or that she doesn't sincerely believe it to be so. Along the same lines, we can understand why an occurrent experience of boredom can lead to a person's interest in doing something she sees as not valuable to do. The descriptions that I offered above, in which a person's akrasia or continence is due to their patterns of interest and boredom, can thus be vindicated.

Moreover, this also makes available an alternative interpretation of people like the narrator of Berryman's poem – one that I think is just as plausible as the first. Rather than simply reading him as an evaluative nihilist who is bored in response to his belief that the world contains nothing of genuine value, we might also interpret him as somebody who sincerely believes there to be good things in the world – *genuinely* great literature, as he suggests – but who finds himself, perhaps to his dismay, uninterested in engaging with such things any longer, in spite of their goodness.

IV. VIRTUES AND VICES OF THE EASILY BORED

I have so far argued that a person's experience of boredom can be justified in a quasi-perceptual manner when the object of her boredom is genuinely boring. One reason why we find somebody like Rawls' grass-counter so unnerving is because he seems to be able to maintain positive interest in something that just isn't *interesting*. Again, perhaps his attention is directed inward, or perhaps he sees something that we don't: but importantly, these thoughts simply reinforce the idea that both reactions are reactions *to* something, and that to be unborable regardless of one's external circumstances suggests either a lack of engagement with, or a distorted perception of, those circumstances.

One might nonetheless wonder whether a person would live a better or more attractive life, or be a better or more appealing person, if she were never bored by, in particular, those things that she recognizes is best for her to do. Wouldn't it be preferable to be fully engaged by these actions and activities, rather than bored by them, and because of one's boredom, at least partly disinclined to do them? To answer these questions, I want to more fully explore what a person's patterns of boredom of interest can say about the sort of character or mind that she has, in general.

A common thought is that a notable predisposition toward boredom always reveals unattractive or regrettable qualities of character and mind. A person who is easily bored may have this tendency because she is deficient in imagination, creativity, curiosity, or perhaps other forms of intelligence; or we might think that though she possesses these capacities, her occurrent boredom is a sign that she is currently failing to exercise them. This latter thought in particular is one that we are likely to have of children who complain of boredom, but it can be extended to adults just as well, accompanying either advice or even criticism. And this criticism can sometimes tend toward a *moral* criticism about one's inability to be fully grateful for what one

has been provided. We might think that she is simply overlooking something that deserves her attention, and her oversight is an indication of a shallow, or ungracious character.¹¹⁷

In some cases, this sort of evaluation seems exactly right. For example, a person may at first find the New Mexican desert maddeningly boring. But perhaps a friend claims that she has just been too dismissive, so she attempts to scrutinize it more carefully the next time she drives through. As she pays closer attention to its details, she begins to notice certain streaks of colour, patterns that have been carved into rock by the wind, and formations of clouds she hadn't noticed before, which a less attentive or perceptive person might never appreciate at all, and which she may not have noticed without taking her friend's suggestion seriously. She might conclude that really, she was mistaken in thinking the landscape so dull; it is in fact, quite interesting. We can also think of activities that may at first appear boring but are revealed to be interesting upon closer inspection and more serious engagement: weeding a garden, bird-watching, or memorizing baseball statistics come to mind.

But although we are right to think that boredom can sometimes indicate a regrettable lack of attentiveness or perception on the part of the person who is bored, a person's particular patterns of boredom and interest can also be explained by a basic difference in taste, rather than a failure to possess or exercise certain mental capacities, or because of ingratitude or shallowness. Another person, preferring the bustle and grit of urban landscapes, might have noticed the very same details of the Nevada desert that her friend now finds interesting, and nonetheless still find

¹¹⁷ For example, in a memorable passage from Marilynne Robinson's *Gilead*, John Ames writes to his son, "I wish I had paid more attention to [water]. My list of regrets may seem unusual, but who can know that they are, really. This is an interesting planet. *It deserves all the attention you can give it*" (p. 28). In another part of Berryman's poem, Henry remarks how his mother had told him never to confess to being bored, as it indicates a lack of "inner resources". And in an episode of the television show *Louie*, Louis CK responds to his four-year-old's complaints of boredom with the following remark: "...you live in a great, big, vast world that you've seen none percent of. Even the inside of your own mind is endless; it goes on forever, inwardly, do you understand? The fact that you're alive is amazing. So you don't get to be bored."

it a bland backdrop only to be endured or ignored rather than enjoyed. Some people simply cannot find the natural world to be of any interest, and it seems possible that in at least some of these cases, it is not because of a lack of attention or shallowness in their values. Some become engrossed by their colleague's fantasy football league; others just cannot bear the thought of participating in this sort of activity. In such cases the suggestion that one pay more attention to what one is bored by may only make things *worse*: perhaps the more some people scrutinize a stand of pine trees, the more dull they find it; the more some people attend to the minutiae of baseball statistics, the more their eyes glaze over.

And importantly, if a person *isn't* bored by what it is that she is doing, or the situation that she finds herself in, it is sometimes because she hasn't scrutinized or attended to her activity or her environment *enough*. For example, a person may be able to watch and enjoy an episode of a boring sitcom that she has seen before only because she is presently exhausted, or mentally depleted. And we should not forget that, as a matter of empirical circumstance, there are also plenty of occasions in which the activity or the environment that a person is faced with is *itself* boring: ill-equipped, or not equipped at all, to provide an engaged and lively human mind with anything of interest. If the world that a person experiences is an interesting place, it is only contingently so; if it is not interesting, heightened attentiveness or perspicacity is likely to only exacerbate one's boredom. And as we found in the case of Rawls' grass counter, if a person were to exhibit sustained positive *interest* in something that is itself boring, we may be tempted to conclude that he is not completely or fully perceiving his activity or his environment.

So, while a person's experience of boredom may sometimes be an indication that he isn't fully aware of the features of his situation, is lacking in creativity, curiosity, gratitude; or hasn't sufficiently reflected on what it is that he is doing, it can also – in cases where it his environment

or activity itself that has very little of interest to offer – indicate that his mind is both discerning, and acutely attuned to the way that the world presently is. His mental capacities may be of good quality, and he may be primed to exercise them; it is the world that yields nothing in return. And if such a person is unable to turn his attention toward something else, being bored by his activity or environment seems to be a perfectly *warranted* response. The advice or injunction to pay more attention to what one initially finds boring in order to alleviate one's boredom, if good or plausible, must take for granted that there is something there, after all, for an attentive and perceptive person to be engaged by.

This brings us to considerations about the *virtues* of mind and character that are connected to a person's susceptibility to boredom. Sometimes a person is easily bored precisely because his mental capacities are of a particularly high quality, or because rather than being incurious, apathetic, or lacking in imagination, he is *especially* curious, inquisitive, or imaginative. Rather than failing to be discriminating or alert, his capacities for discrimination and his alertness may be quite refined. So, in recognizing that boredom can be responsive to features of one's environment or activity, we arrive at a result that is the inverse of the one we began with – both of which seem to me to capture different truths about the relationship between a person's character and mind, and her susceptibility to boredom. A person's disposition may be such that what can keep others fascinated for longer just cannot sustain her interest, because it takes *more* for her mind to be placated. She may be more easily bored than others because her capacities of discernment are so acute, and her mind so quick, that she can speedily observe and absorb the details of something, recall similar instances of that thing, more quickly arrive at the desire for something novel, and so be more susceptible to (and perhaps more averse to) the feeling of boredom when nothing novel avails itself. It is a familiar observation that while some

students are bored in class because unable to grasp the material or unwilling to engage with it, others are bored precisely because of a quick grasp of the material, and a subsequent desire to move on to something new. And we recognize this positive set of characteristics in some people that we greatly admire: Sherlock Holmes's susceptibility to boredom – a feature of his character that Watson describes as a major, defining difference between them – is inseparable from the sharpness and intensity of his mind, and his high level of intellectual energy.¹¹⁸

These different relationships between a one's tendencies toward boredom and the quality of one's character can explain why it is that we are often *ambivalent*, rather than simply critical, in our assessments of people who seem easily bored. We wonder whether there is something more that they could do to more carefully perceive or appreciate their present circumstances, and whether they are being too hasty in finding them uninteresting; we may even disapprove and chastise them when their boredom leads them to do silly or reckless things. But nonetheless, it seems that simultaneously, we can also be sensitive to the possibility that a person's susceptibility to boredom may be evidence of either her (perfectly acceptable) tastes, or even of qualities of mind and character that we appreciate, enjoy, and even admire.

As an example of this, take the reactions that one might have toward the protagonist of Jane Austen's *Emma*. We recognize that Emma Woodhouse's meddling in others' affairs is largely the result of the fact that she is (though she herself doesn't seem to realize this) frequently bored by her life as a caretaker for her aging father. And we are primed by Austen to see her tendency toward boredom as a consequence of certain flaws in her character: her impatience with tasks that take time and persistence to complete, as well as her lack of interest in developing any

¹¹⁸ In "The Singular Experiences of Mr. John Scott Eccles, Sherlock remarks: "...you know how bored I have been since we locked up Colonel Carruthers. My mind is like a racing engine, tearing itself to pieces because it is not connected up with the work for which it was built. Life is commonplace, the papers are sterile...." (p. 224).

of her talents seem to confirm that, being “handsome, clever and rich”, she has gotten away in life without having to develop the qualities of fortitude or persistence.¹¹⁹ Impatient with more acceptable tasks, she becomes idle; her idleness feeds her imagination, and her imagination gets the better of her, making it difficult for her to see people as they really are, rather than as pieces in her matchmaking schemes and invented narratives. Given all of this, we are ready to find her heedless, immature, and irresponsible.

But at the same time, we may also suspect that her scheming and impatience may in fact be the consequences of her possession of rather *good* qualities of mind and character such as cleverness, playfulness, sharpness, and liveliness. We are told, for example, how Emma has from a young age grasped things much more quickly than her older sister; and how with her governess and friend, Miss Taylor, married and moved away, Emma finds herself “in great danger of suffering from intellectual solitude.” Austen continues, “She dearly loved her father, but he was no companion for her. He could not meet her in conversation, rational or playful.”¹²⁰ In other words, Emma has a certain spark or vivacity about her; she’s a firecracker. We sense that she has been placed in circumstances that though perhaps another person (reserved and placid Jane Fairfax, for example) would find sufficiently absorbing, a person with Emma’s qualities of mind would not. Given her qualities of mind and character, she is “in danger”, too, of finding her circumstances boring, and her boredom intolerable – no wonder, then, her tendencies toward mischief.

¹¹⁹ *Emma*, p. 1.

¹²⁰ *ibid*, 6-7.

V. AKRATIC GRASS-COUNTING

So far, I have argued that though there may be forms of boredom that one feels in response to a lack of value in the object of one's boredom, we should nonetheless recognize that sometimes boredom is simply a response to the *boringness* of an activity, and not necessarily its lack of value. This leads to the result that sometimes we are enkratic or akratic because of our experiences of boredom, and interest. And in the last section, I argued that though we often think of the susceptibility to boredom as being revelatory of a vice or some other character flaw, that this is too limited an assessment to capture the variety of ways in which we assess easily bored people. Under certain circumstances a person's boredom may be the result of a simple difference in taste, or her accurate assessment of her circumstances; moreover, her tendencies toward boredom may be inseparable from certain attractive qualities of mind.

I shall now make explicit the results we reach from bringing these two conclusions together. A person's ambivalence toward performing the best action, when explicable by her finding that action boring, as well as the motivations that she has to perform bad actions because of her occurrent experiences of boredom, need not be indications of a defective or unattractive mind or character. On the contrary, the positive qualities of character that a person's susceptibility to boredom can be indicative of may also be present – and may even be properly conditioned in a way I'll elaborate on – in those who are bored by what they know would be best to do.

To illustrate this, let's consider two alternative cases of grass counting. Wholehearted Harriett and Enkratic Edith, both graduate students in biology, are counting blades of grass for a research project. So, unlike a grass counter who simply – *somehow* – finds this task engrossing in itself, Harriett and Edith are engaging in this activity because it is instrumentally valuable for

them to do so. Let's imagine also that in order to do this task successfully, neither can simultaneously distract themselves from what they are doing, in the way that a person might listen to the radio or daydream in order to alleviate the boredom they might otherwise experience. If they are to count blades of grass successfully, they must exert their full attention to what it is that they are doing.

Edith, in finding this task exceedingly boring, experiences some motivational resistance toward performing it. Though she recognizes that she has decisive reason to continue counting blades of grass, her boredom with grass counting renders her less than whole-hearted about doing it. While she loves her work in general, she cannot love this aspect of it. Half an hour into the day, she finds that she must muster self-control in order to resist the temptation to get up and stretch her legs, get another cup of coffee, or check her email on her phone, all the while recognizing that she has little or no reason to do these things. After all, it is not as though her legs really are cramped, or that she really wants a cup of coffee, or that she needs to check her email; and she is fully aware that these actions will just delay the work that will have to be done anyway.

Harriett, like Edith, knows that the best thing for her to do is to count blades of grass, but unlike Edith, she is *not* motivationally divided. Rather than being strong-willed in resisting the temptation to do something else in order to avoid this boring task, she isn't tempted, to begin with. She does the best thing, emotionally and motivationally undividedly, for the right reasons. She is a *wholehearted* grass counter, lacking the desires that Enkratic Edith has to stop counting and do something else instead.¹²¹

¹²¹ I want to address a worry that one might have at this point. Though in this example, what is best for both Edith and Harriett to do is *instrumentally* best for them to do, we shouldn't be misled into thinking that it is only actions with instrumental value that can be both boring and good to do, as though good actions and things are more susceptible to being boring when their goodness is purely instrumental. There

At least initially, I think we are likely to find Wholehearted Harriett's psychology puzzling in a way that we don't find Enkratic Edith's, whose experience is likely to be more familiar to us. How exactly does Harriett experience this activity? In trying to imagine the virtuous psychology in a charitable manner – one that yields to us a clearly attractive human psychology – we will want to accommodate that she perceives her circumstances accurately, so she must be able to appreciate that grass counting is, after all, a boring activity. And as discussed earlier, given her appreciation of the boringness of what she is doing, it would be bizarre if she were utterly *engrossed* by the task; this would suggest at least some sort of insensitivity to what she is up to. So her wholeheartedness, in this sort of case, should not be understood as a matter of taking pleasure or being positively enthused by what it is that she is doing.

So consider the supplemented view that I began with, which allows that unity or wholeheartedness doesn't demand positive enthusiasm or pleasure. Rather, the virtuous agent

are intrinsically good activities, as well as intrinsically good things, that are *also* boring, which can be experienced *as boring* even by those who are interested in the kind of goodness that is in question. For example, it is a familiar remark of philosophers that Sidgwick's *Methods of Ethics* is both good *and* boring. It has also been suggested, plausibly, that these two features are inseparable from one another: that some of the properties that make the *Methods* good are the very same properties that make it boring. Here is C.D. Broad's description in *Five Types of Ethical Theory*, of his experience of reading the *Methods* (which he also describes, at the top of the chapter in which this quote appears, as "on the whole the best treatise on moral theory that has ever been written"):

[Sidgwick's] style is heavy and involved, and he seldom allowed that strong sense of humour, which is said to have made him a delightful conversationalist, to relieve the uniform dull dignity of his writing. He incessantly refines, qualifies, raises objections, answers them, and then finds further objections to the answer. Each of these objections, rebuttals, rejoinders, and surrejoinders is in itself admirable, and does infinite credit to the acuteness and candour of the author. But the reader is apt to become impatient; to lose the thread of the argument; and to rise from his desk finding that he has read a great deal with constant admiration and now remembers little or nothing. (p. 144)

My point here isn't to bring up the virtues and vices of pieces and styles of philosophical writing (one might point out, after all, that philosophy *can* be made worse off by being boring). Instead, it is just to highlight the possibility of something's being *both* intrinsically good and boring. Again, that something good makes no promise of also being interesting, and that something bad makes no promise of also being uninteresting is precisely why it can take a concerted effort to attentively read Sidgwick's writing on the one hand, and to refrain from gawking at train wrecks, both literal and figurative, on the other.

appreciates that grass counting is boring, but the consideration provided by the boringness of grass-counting plays no other role in her response to it. She does not positively enjoy this activity, but she does not find it *painfully* boring, or boring enough that it would disincline her from doing it.

While I think that this proposal is initially difficult to fully imagine, there are resources that we can rely on to make her psychology more vivid. For example, we might think that by focusing on the *nobility* or *fineness* of her ultimate goal – advancing scientific knowledge – she is able to recognize that what she is doing is boring, but have that boredom be completely “silenced” so that it does not at all impact her motivations.¹²² Perhaps the most plausible rendering of her psychology is that she experiences a mental state akin to a meditative tranquility, which is capable of resisting the experience of grass counting boring in any way that would leave her anything less than wholehearted about doing so, without compromising her perception of her circumstances.¹²³

No doubt being able to maintain this inner calm is an attractive capacity – and one that we are right to admire. And it is likely supported by a number of virtues, such as patience, calmness, or a kind of stoicism. The question that remains is whether it is a capacity possessed by those with a mental life clearly *more* attractive than or preferable to the enkratic’s, whose will is divided by her aversion to performing boring tasks. Though it is true that Edith must struggle in a way that Harriett does not in order to do what she knows would be best to do, though her experience of this activity will be subjectively unpleasant in a way that Harriett’s may not be, and though there is the risk that she might opt to do something else in order to avoid her boring

¹²² This psychological capacity is highlighted by McDowell (1978).

¹²³ In the psychological literature this sort of state is referred to as “flow”, or “zone”.

task, I propose that the fact that she is motivationally divided rather than virtuous in regards to this activity need not be any indication of a character or mind that is worse off. Moreover, she may possess certain positive qualities that Harriett does not.

First, Edith's boredom with grass counting – and her robust experience of it as such – is a perfectly *worthy* response for her to have, given what the activity of grass counting is actually like. Second, as I've suggested, that a person is particularly susceptible to boredom may indicate that her mind is particularly lively or discerning, or that she may possess certain forms of intelligence that make it difficult for her to be patient and persistent with dull tasks.

This second observation holds, I propose, even in cases where the action that the agent finds boring is also the action that would be best for her to perform. In fact, that Edith *knows* what would be best to do¹²⁴, and nonetheless still desires to either stop, or do something else entirely, seems to be important to the thought that her experience of boredom may be indicative of these positive qualities of mind. This is why, as I suggested, the enkratic's susceptibility to boredom is properly conditioned in a way that renders it unlike the vices associated with boredom. Her boredom is not best explained, for example, by a failure to appreciate the value of something – the way in which we might explain and criticize the boredom of a child, or a misguided or shallow adult. Again, recall our ambivalence about Emma's susceptibility to boredom: it isn't immediately obvious whether her boredom is brought on by a failure to fully appreciate what the world has to offer, perhaps because she has mistaken beliefs about the value

¹²⁴ Perhaps this is something that a defender of the virtuous agent will push back on: Edith couldn't possibly fully understand the value of what she is doing – and this is precisely why she feels resistance to grass-counting. If only she really *knew* how valuable scientific inquiry was, she'd be wholehearted about doing it. There are several difficulties with this line of thought. It's hard to see exactly what it is that Edith lacks, and Harriett has (*vis a vis* evaluative knowledge). And this would make the conditions for knowledge of value implausibly high – in a way that seems plainly unrealistic. But most importantly for my purposes, it is hard to see, even if we grant that Edith has some “lower grade” of knowledge, how this would actually help settle the question of whose psychology is more preferable, or attractive.

of the activities that she finds interesting and those she finds boring, or whether it is just a result of her positive qualities of mind placed in limiting circumstances. But if Emma really does know the value of such things – if we don’t think that she’s simply being shallow or imperceptive – we’re more likely to see her in this second, more positive way.¹²⁵

However, one might be worried by my proposal that Edith’s disharmony may be something worth preserving, given that by hypothesis, what she is bored by is what she has most reason to do. Her motivations, we are to imagine, are out of proportion with or fail to cohere with what would be best, and so are necessarily partly in favour of doing something that would be worse: and in this circumstance in particular, it’s simply to just *not* do what she should do. Surely the virtuous person experiences boredom in a way that disinclines her from certain actions as well – just not when she has decisive reason not to. And furthermore, we might recall at this point boredom’s tendency to give rise to motivations to not only perform actions that aren’t the best, but actions that are positively bad, pointless, silly, or stupid. So being easily susceptible to boredom, which I have been highlighting as something potentially *positive*, is the very same disposition that, when possessed by the bored enkratic, necessarily involves a person’s being motivated to perform actions that are either not the best she could perform, and are sometimes positively bad to perform. And if such a person were to actually perform such an action

¹²⁵ This emphasis on the importance of evaluative knowledge as a condition on virtue renders the qualities I’ve focused on psychologically importantly similar to those that a standard Aristotelian account of the virtues would defend, and deeper than the qualities of character that count as “virtues” on certain more minimalistic theories of virtue, such as Hume’s. While I think Hume is right to turn our attention to the wide variety of qualities that we appreciate in ourselves and one another, I think that Aristotelian accounts are right to emphasize that there is an important difference between the qualities of character that would count as virtues on Hume’s account, and thicker qualities of character that reveal to us more about what the virtuous agent’s mental life, and capacities, are like.

akratically, citing her boredom would not serve as any *justification* for her doing so – it would only serve as an explanation.

Nonetheless, given the connection that this susceptibility has to certain other characteristics of mind that we admire and appreciate, it's not obvious to me that we would wish for others, or ourselves, to be less vulnerable to boredom – even when that vulnerability leaves us bored by things we also recognize are valuable and good. While I think that we are right to be ambivalent about the enkratic's susceptibility to boredom, our ambivalence rather than our outright disapproval marks that we nonetheless would not be so quick to rid ourselves and one another of it as we might be encouraged to by an antecedent commitment to the thought that the virtuous agent's psychology is clearly and obviously preferable to the enkratic's. Holmes' tendency to become easily bored is inseparable from other excellent qualities of mind and character that we positively admire. And so, we should appreciate his susceptibility to boredom even if – to Watson's concern – it is the very same tendency that nourishes his cocaine habit when he is left with nothing interesting to do.¹²⁶ He wouldn't be who he is, with the kind of attractive and appealing character he has, without it. And because of this, I think it would be natural to see his tendency toward actually performing akratic behaviour in a more positive light, overall: one that presents it as an essential and charming feature of a vibrant and brilliant soul.

If this is right, then we must reconsider whether it would be obviously better, preferable, or more ideal should someone like Edith become more like Harriett, Holmes to become more like Watson, Emma to be like Jane, and in general, whether bored enkratics would possess more

¹²⁶ As Watson reports in "A Scandal in Bohemia": "My own complete happiness, and the home-centered interests which rise up around the man who first finds himself master of his own establishment, were sufficient to absorb all my attention, while Holmes... remained in our lodgings in Baker Street, buried among his old books, and alternating from week to week between cocaine and ambition, the drowsiness of the drug, and the fierce energy of his own keen nature" (p. 5).

excellent characters if only they were more wholehearted, and less bored, by what they know they should do. Given that there need be not be anything defective about the bored enkritic's perception of the situation in these cases, and given that her susceptibility to boredom may be inseparable from, because the direct result of, her possession of certain positive and attractive qualities of character, it may in fact be something that we would, I think rightly, regret.¹²⁷

VI. CONCLUSION

If what I've argued here is right, we should re-examine the extent to which we've privileged the psychology of the wholehearted over those who are more susceptible to *enkrasia* or *akrasia*. Sometimes the disharmony we experience is simply the result of an intelligent response to the features of one's environment. Again, this is not to suggest that there isn't something attractive or desirable about a person's ability to be wholeheartedly committed to the best action, no matter how boring, or otherwise unattractive, that action is. If it is psychologically possible for a person to be able to fully appreciate the boringness of something while still remaining wholehearted about it as long as one also fully appreciates the goodness of what one is doing, then perhaps this is a capacity possessed and reliably exercised by the more patient, steadfast, and stoic among us – one worth admiring and respecting for its own sake. Rather than denigrate this capacity and the virtues it underlies or is supported by, my hope is to encourage a

¹²⁷ One might argue that while the positive qualities of character that I have highlighted may be *important*, that should they be subversive to a person's *moral character*, we have good reason to disvalue or disregard them; and perhaps this is why the virtuous person's character is more attractive. Here, I would emphasize that these qualities need not be incompatible with a person's moral virtue, but that furthermore, even when they are, we might remind ourselves of just one point argued for by Susan Wolf in "Moral Saints": that upon reflection on the sorts of people that we love, and the relationships we cherish, the qualities of character that we find of deep importance are not identical to, and may even conflict with, the moral virtues. It is important to note, however, that while Wolf focuses on a person who is maximally morally good, I have focused on the question of the importance of wholeheartedness to the virtues, whether or not those virtues are best understood as "moral" in nature.

wider investigative scope, and a pluralism when it comes to the philosophical project of delineating the positive qualities of character and virtues: a scope that includes, and a pluralism that recognizes and celebrates, those excellences that depend on or are inseparable from a person's emotional ambivalence or aversion toward what she knows would be best to do.

Fearsome Goods

I. INTRODUCTION

I will here discuss a dogma that has come to shape much philosophical discussion about the nature of value: a conception of value which reinforces, or relies upon, the basic philosophical approach toward our motivational and emotional responses that I have been resisting over the last three chapters. It can be expressed by two parallel claims: for something to be good is just for it to be something that it would be fitting or appropriate to have a positive attitude, such as desire or admiration, toward; likewise, for something to be bad is just for it to be something that it would be fitting or appropriate to have a negative attitude, such as hatred or antipathy, toward. This connection between value and our appropriately held emotional and motivational responses has been taken by many to be the most promising way to think about the nature of goodness and badness. It is, after all, more illuminating than the intuitionist's conclusion that goodness is simply unanalyzable.

Claims of this basic form serve as the foundation of a family of views that have recently been dubbed *fitting-attitude theories of value*.¹²⁸ Such views hope to reduce and explain value in

¹²⁸ Contemporary work on “fitting” attitudes includes D’Arms & Jacobson (2000), Zimmerman (2011), Rabinowicz and Rønnow-Rasmussen (2004), and Anderson (1993). Fitting-attitude theories of value are often attributed to historical figures such as Brentano, Ewing, and Sidgwick. Scanlon (1998) also endorses a version of a “fitting attitude” account, embedded within his “buck-passing” view of value. According to a buck-passing view, to say that *x* is valuable is to say that *x* has some other “natural” property that gives us reason to have some positive attitude (typically one with practical implications) toward *x* (p. 96). The fitting-attitude claim – just about the connection

terms of appropriately held human attitudes.¹²⁹ But one need not share these theoretical ambitions in order to find the original analysis intuitive: realists, too, may be committed to this intimacy between value and our appropriately held emotions and motivations without feeling the need to be able to reduce one to the other.¹³⁰

These claims are taken as uncontroversial in other debates, as well: both non-cognitivists and cognitivists about evaluative or normative judgment take pains to accommodate or establish this intimacy; others have defended the view that should a person desire something (that is, should he possess a paradigmatically “positive” attitude toward it), he must be doing so under the auspices of its goodness. And even those who defend the possibility of a person’s motivations and emotions diverging from her evaluations are quick to point out that there are pressures on her to eliminate this tension or division – and this last observation is precisely what might be best captured by the notion of the “fit” or “appropriateness” of one’s responses.¹³¹ These psychological and normative claims are consonant with, and supportive of, the thought that the structure of value is perfectly isomorphic with the objects of our positive and negative attitudes, correctly held.

There is, no doubt, something that this basic picture gets right. There are many contexts in which to state that something is good or bad is just to recommend, or expect the formation of a

between something’s being good and having a reason to have a positive attitude – can be separated from the claim about the “naturalness” of the reason-giving property.

¹²⁹ For example, in “Practical Reason and Incommensurable Goods”, Elizabeth Anderson endorses and attributes to Kant what she calls a “pragmatic theory of value”, according to which “value judgments are constructions of practical reason that guide our reasoning about what to do and what to care about... instead of saying that it is rational to value something because it is good, pragmatism says that it is good because it is rational for us to value it. Claims about what is rational to value determine claims of value” (p. 92).

¹³⁰ Moore, for example, writes, “The desirable means simply what ought to be desired or deserves to be desired; just as the detestable means not what can be but what ought to be detested...” (*Principia Ethica*, pp. 66-7).

¹³¹ Among others: Smith (1987), Stocker (2004), Hurka (2003), Anderson (1993), Sussman (2009).

certain attitude toward it; in many cases, to ask whether something is good or bad is just to ask how one should feel, or what to do about it. Furthermore, should a person fail to have a certain emotional or motivational response toward something we think of as good or bad, we may think that they do not have the same perceptual or cognitive grasp of the thing that we do. We often urge one another to look again, and in so doing, invite a new response – one that resonates with the value that we see, and the emotions that we feel.

These observations can lead to the impression that the relationship between evaluations and our appropriately held emotions is a truism that any successful philosophical view about value or human psychology must accommodate. Nonetheless, even though these claims are able to capture much of our thought and discussion about value, it is important that we recognize their limitations. Unchecked, these claims threaten to distort our understandings of our own attitudes and the reactions we have to one another in light of those attitudes. Moreover, they place objectionable limits on the imaginative possibilities of goodness and badness, closing off the possibility of asking certain important philosophical questions for the sake of theoretical simplicity.

II. THE CLAIMS

I will focus my discussion on two claims. For ease of discussion, I present them as two bi-conditionals:

x is good iff x is a fitting object of a positive or pro-attitude

x is bad iff x is a fitting object of a negative or con-attitude

Let me begin by briefly elaborating on a few preliminary clarifications about how best to interpret these claims. First, let's grant that the basic thought behind these claims is capable of

accommodating the distinctions between, for example, intrinsic and instrumental value, as well as personal and impersonal value. It should be no difficulty for a view that supports these claims that a person may feel affection toward a ticket stub from the first concert she attended, or toward the particular shade of blue of her childhood home, while not thinking these things good in any *impersonal* sense. On a full elaboration of these claims, she need only think that that stub, or that colour, are valuable *to her*. Second, these claims need not be committed to some *particular* positive attitude or negative attitude that it is appropriate to have toward the good and the bad. For example, while some fitting-attitude theorists refer to *desire* or *love* when considering the attitudes fitting to have toward goodness, we should read these claims as inclusive of other positive attitudes such as admiration, respect, or interest¹³², as well as attitudes connected to practical stances such as preservation, promotion, and protection.¹³³ Third, such claims should be understood as claims about value, understood broadly – not just, for example, *moral* value.

But important questions remain that will take a bit more discussion to demonstrate the relevance of. There are at least two major ambiguities that must be addressed when considering whether these really are basic truths about goodness and badness. First, we must determine what sense of “fit” such claims are committed to. Second, we must determine whether there is one coherent way to distinguish between those of our attitudes that are “positive”, and those that are “negative”, such that the claims that we fittingly feel the “positive” ones toward objects of goodness and “negative” ones toward objects of badness are not merely stipulative.

¹³² Brentano uses “love”, while Sidgwick uses “desire”. Elizabeth Anderson and T.M. Scanlon emphasize the importance of including a wide variety of value-centered attitudes – both see this as the only way to accommodate a wide enough pluralism about value. I do not think that that argument is sound, but I will not discuss it here.

¹³³ Though Scanlon does discuss some emotional responses that could in principle be felt independently of action, he places a heavy emphasis on attitudes more closely connected to action.

First Ambiguity: “Fit”?

The language that philosophers have used when characterizing this fit ranges widely: some use the prescriptive “ought”, “should”, or “right”, but others use the weaker, “appropriate”, or even the weakest, “make sense”.¹³⁴ This casual slide masks an issue that must be addressed, as different senses of “fit” will yield claims that are distinct in content and plausibility.

To illustrate the importance of making sure we know what “fit” we’re after, consider, for example, this interpretation: that those things that it is *morally appropriate* to have positive attitudes toward are those things that are good. As Justin D’Arms and Daniel Jacobson have argued, this thought involves a conflation of a moral sense of fit with a non-moral, but nonetheless important, sense of fit.¹³⁵ To illustrate, they note that Elizabeth Anderson shifts between these senses when she writes:

A person may laugh at a racist joke, but be embarrassed by her laughter. Her embarrassment reflects a judgment that her amusement was not an appropriate response to the joke. The joke was not genuinely good or funny: it did not merit laughter.¹³⁶

As D’Arms and Jacobson point out in their response to Anderson, though there are no doubt connections¹³⁷ between them, standards of humour are not identical to moral standards.

¹³⁴ Sidgwick, Moore, and Brentano use “ought”; Scanlon uses the language of having a reason to feel or do something; Anderson varies between “ought”, “appropriate” and “makes sense”.

¹³⁵ D’Arms, J, and D Jacobson. “The Moralistic Fallacy.”

¹³⁶ *Value in Ethics and Economics*, p. 2.

¹³⁷ Some comedians, for example, think that comedy should “punch up, not down”. Moreover, there do seem to be jokes that are not funny precisely because of how morally bad they are, or because of the morally bad attitude that underlies them. But the important point, consistent with these, is that there are cases in which the norms of humour and the norms of morality may genuinely conflict with one another.

There are things that are genuinely funny and so appropriate, in some sense, to laugh at, even though they may not be *morally* appropriate to laugh at.

One general lesson we can draw from their discussion of Anderson that is germane to the topic here is that although it may seem as if we can simply rely on some basic, intuitive sense of appropriateness when considering what we mean for a positive attitude to “fit” an object of positive value, this fit has a particular sort of character. We should not expect this kind of appropriateness to have the same character as moral appropriateness, or prudential appropriateness: this would leave us with a realm of value that is stunted by the demands of morality or prudence, in the way that it seems as though Anderson’s discussion of a racist joke has conflated the funny with the morally good. Furthermore, the sense of fit we are after doesn’t seem to have the force of a *prescription* or a *command*.¹³⁸ While I might think it fitting that you find funny a comedian that I think is funny, it would take quite unusual circumstances to think that you *ought* to find him funny, where we take this “ought” to signify anything like an obligation or a claim about your rationality. As D’Arms and Jacobson put the point, “...there is a crucial distinction between the question of whether some emotion is the right way to feel [i.e. morally right, prudentially right, all-things-considered right], and whether that feeling ‘gets it right’”.¹³⁹

D’Arms and Jacobson are exactly right to recognize this distinction – one that takes into consideration whether or not an attitude is one that it is appropriate to feel given some independent (moral or prudential or all-things-considered) requirement or consideration, and

¹³⁸ Perhaps an implicit commitment to a form of consequentialism has made it plausible to many fitting-attitude theorists that the “ought” here really is a moral ought. See, for example, Zimmerman (2011). If this is the case, then the arguments I present here should equally work as arguments against the form of consequentialism informing these analyses of value.

¹³⁹ “The Moralistic Fallacy”, p. 66, my emphasis.

whether it is appropriate to feel simply because it “gets it right”. Nonetheless, the idea of a feeling “getting it right” or wrong is itself ambiguous. Return again to the example of humour. We all know very young children who will laugh at anything having to do with the toilet. There is a genuine and important sense in which the statements or situations that could get these kids going are *not* funny. Their laughter, in a genuine and important sense, is unfitting.

But there is another sense in which their laughter *is* fitting. Consider, for example, a child’s laughter at the words *poop deck*, used in nautical contexts. Compare your reaction to this with the reaction you would have toward a person who laughs at the fact that fingers are attached to hands, or that sometimes it rains. A person who laughed at these latter things isn’t simply laughing at something *unfunny*: he’s laughing at something we cannot find it *intelligible* to be amused by. This is unlike the way that toilet humour or staged pratfalls or terrible puns *are* intelligible for a person to laugh at without being, we might think, *actually* funny.¹⁴⁰ Our reaction to somebody who laughs at the first set of things is unlike the reaction that we have to somebody who laughs at the second. While we might scold or criticize a child for his laughter (or rather, his parents might), we still *understand* why he is laughing.¹⁴¹ In contrast, we will be completely baffled by this other person: so much so that we may not know *how* to react to him, thinking instead that he is deranged or delusional.¹⁴² As I discussed in Chapter 3, this sort of reaction, and

¹⁴⁰ To echo points that I made in chapter 3, the terminology we might use to mark this distinction is between the *afunny*, and things that are genuinely funny *and* things that are unfunny. The *afunny* is reflected in our responses to the different ways in which the responses are “unfitting”. There are some comedians that explore unfunniness, or anti-humour: Andy Kauffman, or Harris Wittels, for example.

¹⁴¹ One might worry here that the sort of understanding we have of my nephew is too thin to mark the distinction I am after here: for example, some human behaviour might strike us as unintelligible, but we’ve simply gotten used to it. If this were right, then either things are *afunny*, or simply funny – things that are “unfunny” are either one or the other. I want to resist this collapse, however: it doesn’t seem to truly capture the thought that there are things that are distinctively *unfunny* – and that these differ from both things that are just “not funny enough” and things that are *afunny*.

¹⁴² The director David Lynch is particularly good at creating this experience. In his short film, *Rabbits*, a laugh track is added to a disjointed and cryptic conversation among a family of anthropomorphized rabbits. Apropos nothing,

the fact that it is distinct from the sort of reaction we have the child who laughs at “poop deck”, for example, is important for us to recognize – I’ll come back to this point at the end of this section.

Given this, there are two ways in which an attitude can “make sense” to have: there is the sense of fit in which laughing at a *good* joke is fitting and laughing at “poop deck” is not fitting, and another sense of fit – having to do with the *intelligibility* of the response – in which it is, for example, fitting to laugh at the words “poop deck”, but isn’t fitting to laugh at an unremarkable patch of pavement or a certain shade of yellow. We need to recognize that these two senses come apart from one another, otherwise we may illicitly rule out as simply unintelligible a large scope of reactions that *do* “make sense” for a person to have, even though they may not track the sort of value or disvalue taken to be central to that reaction. But insofar as we are trying to give an account of what is actually good (or actually funny), we cannot be interested in this second sense – it would declare too many things of actual value when they are not. So while other senses of “fit” like a moral fit or prudential fit seem too restrictive, there is at least one sense that is too permissive.

Is there anything substantive we can we say about this goldilocks sense of fit we are after? What do we mean when we say that a person’s laughter is a fitting response to a good joke, but not a fitting response, in this same sense, to a bad one? Perhaps in this second case, the person has a mistaken belief: his laughter evinces his belief that what he laughs at is funny, when it isn’t. But imagine that as a child grows up, he eventually learns and endorses the thought that simply saying the words “poop deck” is *not* funny. Nonetheless, imagine that as an adult he might *still* find himself, from time to time, laughing at their use. He doesn’t have a false belief,

members of the family state things such as, “*Don’t forget that today is Friday*”, or “*What time is it?*” and the “studio audience” laughs uproariously. The whole thing is just really unsettling.

his response doesn't seem insane, nor does he seem in any obvious sense irrational. A similar point can be made for other cases in which a person forms positive attitudes toward things that they know are bad, or negative attitudes toward things that they know are good: while there is a sense in which this reaction may be "unfitting", it isn't a sense of unfittingness that seems easily identifiable with a form of irrationality or unintelligibility. That a person loves or desires to watch what she believes is a bad television show, or hates a movie she believes is good need be no mark against her *rationality*.

We might begin to wonder whether the difficulty in determining the sense of fit invoked by the claims under discussion should be enough to raise doubt about whether they can be interpreted as both contentful and plausibly true. At the same time, perhaps this line of questioning is bordering dangerously on the pedantic or the obtuse. After all, there seems to be a commonsense way in which, when something isn't funny, it isn't fitting to laugh at it in the way that it is fitting to laugh at things that really are funny. It is the same sense of fit, we might think, that we can use in cases in which a person is scared of something that is actually dangerous or struck by awe by something that is actually awesome. While we don't think people are *obligated* to be afraid of things that are genuinely dangerous, we nonetheless judge that a person's felt fear in response to a dangerous thing is "fitting" simply given the fact that the thing is actually dangerous – no other sort of consideration needs to be used to justify her response. I'll call this kind of fit the *fit of veridicality*.¹⁴³ So we can now interpret the bi-conditionals I started with, keeping in mind that the relevant sort of fit is this fit of veridicality.

¹⁴³ This result poses a difficulty for fitting-attitude theories of value that, on pain of circularity, must avoid relying on evaluative properties when providing content to the sense of fit they rely on in their analyses. As Ross writes in *The Foundation of Ethics* of the attempt to provide an analysis of the attitude of admiration:

Admiration is not a mere emotion; it is an emotion accompanied by the thought that that which is admired is good. And if we ask on what ground a thing is worthy of being thought to be good, only one answer is possible, namely that it is good. It would be absurd to say that a thing is good only in the sense that it is

Second Ambiguity: “Positive” and “Negative”?

Consider now the second ambiguity in how to interpret the two bi-conditionals I started this section with: how should we determine which of our attitudes count as “positive”, and which count as “negative”? For ease of discussion, it would be best to focus on the right-to-left conditionals of the two bi-conditionals I started with, focusing just on the two conditional claims that if x is a fitting object of a positive or pro-attitude, x is good; and if x is a fitting object of a negative or con-attitude, x is bad.

In determining how to understand what it is for an attitude to be “positive” or “negative”, perhaps we should focus on attitudes that are either phenomenologically “positive”¹⁴⁴ in some way, or are closely connected to “positive” practical stances, such as pursuit, promotion, or protection. Natural candidates for positive attitudes might be desire, admiration, respect, love, and fascination; candidates for negative attitudes might be aversion, fear, hatred, disrespect and disinterest.

The worry is that these emotions and motivations can be regimented such that those that are “positive” and “negative” in the sense specified are *also* those attitudes that are best understood as what I will call “value-centric”: that is, veridical when directed toward goodness and badness, respectively.

worthy of being thought to be good, for our definition of ‘good’ would then include the very word ‘good’ which we were seeking to define (pp. 278-279).

It is not clear how many contemporary philosophers have (or should have) the sort of ambition of reduction, for which this circularity would be a problem. Moreover, even if we cannot vindicate *those* theories, we might nonetheless arrive at an informative illumination of the connection between value and our responses: goodness and badness is whatever our positive and negative attitudes pick out when those attitudes are *veridical*.

¹⁴⁴ I hope the idea of being phenomenologically “positive” is intuitive enough, without attaching this either to beliefs about goodness, or with pleasure. For example, *respecting* an object might be phenomenologically “positive” in this way, without being pleasant. Both Scanlon and Anderson seem to understand positive and negative here either in terms of phenomenology, or connection to “positive” actions, but neither elaborate on this point.

Consider again, for example, the contrasting attitudes of interest and boredom. On this understanding of which attitudes count as “positive”, being interested in something or finding it fascinating would be a paradigm of a positive attitude. We might think that it just *feels* good to be interested in something: this is especially salient in comparison to the way that boredom feels. Furthermore, interest bears a connection to certain “positive” actions: we’re likely to seek out and attend to things that we find interesting, and perhaps even promote or protect them. Conversely, we’re disposed to avoid things that we find boring, and it is difficult to attend to them.

But why think that interest is best understood as veridical when directed toward things of *value*, or the good aspects of an object, rather than other features it may have? A person’s positive interest in something may simply be a response to the fact that the thing in question is *interesting*, or *novel*, not that it is *good*. And likewise, as I argued in chapter 3, a person’s boredom in something may simply be a response to the fact that the thing in question is *boring*, or *repetitive*, not that it is *bad*. While it is not a genuinely Chinese curse to wish on somebody that he or she live in interesting times, it is a genuine curse.

Other “positive” and “negative” attitudes have a similar structure, as well. As I argued in Chapter 1, desire – taken to be a paradigmatically positive attitude – need not be understood as necessarily possessed under the guise of the good. There are objects of desire that are not just *intelligible* to desire, but we might think *veridical* for human beings to desire, but which need not be good: things that are merely attractive, or fun, or exciting, or convenient, for example, but also things like domination or violence. By taking our positive and negative attitudes to be value-centric in general, we distort those that are not, and fail to see the sorts of considerations they may actually be sensitive to, instead.

To further elaborate on this point, I want now to discuss a familiar objection that has been raised to fitting attitude theories of value. This objection runs as follows: notice that there can be good reasons for a person to have a certain attitude toward something that don't seem to make reference to the qualities of that object. So, for example, one might have reason to laugh at an unfunny joke made by one's boss, perhaps knowing that he eventually fires every one who works for him except for toadies and doormats. And so though it is fitting for his employees to laugh at a joke that he makes, it isn't fitting because of the funniness of the joke – it's fitting because his employees have prudential reason to laugh at it.

In response to this worry, some fitting attitude theorists have relied on a distinction between what Derek Parfit has called “state-given” and “object-given” reasons.¹⁴⁵ The reasons that fitting-attitude theorists are interested in are reasons that one has given the properties of the object in question, and not given the other considerations that speak in favour of one possessing the attitude, or mental state, in question. A successful analysis will focus on the fact that even though one might have a prudential reason to laugh at one's boss' joke, this is the “wrong” reason for a fitting-attitude theory because it is not “object-given”.¹⁴⁶

I take no stand on whether this response to the “wrong kinds of reasons problem” is successful independently of the point I am making here, or whether it simply begs the question. Nonetheless, it is a helpful way to frame the point I am making here: that there are features or properties of objects that need not reflect the goodness or badness of those objects, but which fully explain the positive and negative responses we have toward them. Consider the following case:

¹⁴⁵ *On What Matters*, Chapter 2.

¹⁴⁶ For a recent elaboration of how this distinction might be used by fitting-attitude theorists to defend their position, see Jonas Olsen, “Buck-Passing and the Wrong Kind of Reasons”.

Jones Jr. learns that his father, Jones Sr., wants to come by his apartment this afternoon and have a heart to heart with him. When Jones Jr. was a child, Jones Sr. seemed to have only two modes of engage: he was either cold and withholding, or angry and frustrated. In the last few decades, Jones Sr. has sought treatment and therapy, and has come to see the damage that he did to Jones Jr., and wants to apologize. Jones Jr., being a particularly forgiving person (and also having gone through some therapy himself), believes that at this point in time, it would be good for them to have this conversation. It would be good for him, for his father, and for their relationship – he believes that his father has changed, and he is ready to forgive him. As the hour of Jones Sr.’s visit approaches, Jones Jr. finds himself growing more and more anxious and distressed by the thought of conversing with his father. He experiences the desire to call the whole thing off.

Jones Jr.’s negative emotions need not be justified in terms of non-object-given reasons, in the way that a person’s laughter at his boss’s joke might be. Instead, his reaction fits this conversation with his father because of salient features of that object. It will no doubt be difficult and awkward, painful – and these are properties of it that not just explain, but justify why Jones Jr. is so averse to having this conversation. His negative reactions are “getting it right”, given those features. But it does not seem right to conclude from these observations about the fittingness of these negative responses that this conversation would be something *bad* – and we do not need to think that Jones Jr. believes that it is bad, or that these features *make* it bad, in any way, in order to understand his reaction. In fact, it seems coherent and sensible that Jones Jr. thinks that having this conversation would be good, in spite of his also having these justified negative reactions toward it.

On the proposal that I am offering, the properties that it is fitting to have different attitudes toward are heterogeneous: distinct from one another, and in principle distinct from the goodness or badness of the object in question. Somebody’s anxiety toward a situation may be fitting simply because the situation is one that involves a serious or life-altering change, for example, not because it is or is believed to be *bad*. Think, for example, of break-ups or divorces between people who know that it would be best for them to call things off and what little comfort they might get from staying together isn’t anything that amounts to something valuable about

doing so. Likewise, somebody's annoyance with something may be fitting simply because that thing is disturbing a pleasant peacefulness, not because it is bad. Think, for example, of somebody who is woken up every morning by his cat pawing at his face. He need not be annoyed because he believes or perceives this early morning greeting as something *bad* – indeed, he might look forward to this daily annoyance, and reminisce about it when he reflects on what he loves about his cat. To summarize the deeper point that underlies the possibility of these examples: to think that positive attitudes, understood as phenomenologically pleasant or positive, are themselves built to pick up on *goodness* must be assuming a closer connection between goodness and pleasure or action than I think is justified.¹⁴⁷

Two responses seem natural at this point. One might claim that the properties that I am referring to just *are*, or perhaps supervene with, properties of goodness and badness. Something's being interesting, for example, just *is* a way for it to be good; something's being boring or frightening just is a way for it to be bad. Or consider this line of thought: those things that it is fitting to hate are just those things that are *worth* hating. And only genuinely bad things are worth hating. And this is why hatred, being a negative attitude, is properly responsive to things that are bad.

Though there are many circumstances in which claims of this sort seem cogent, we should be careful not to generalize these observations. This view about the connection between value and these other properties is controversial, and needs to be argued for, not simply settled or presupposed by the claims under discussion here. And here's just one reason why we might resist such conclusions: whether or not a thing's boringness, or difficulty, or painfulness constitutes a bad feature of it presumably depends on the particularities of the object in question, and the other

¹⁴⁷ This point will get reinforced in the next section. But it is a theme of the dissertation that we should think more seriously about the relationship between goodness and our positive responses – we have assumed that the relationship is far more simple than it is.

standards that we might use to evaluate objects of that kind. Aesthetic objects like paintings and films are often frightening or disturbing, when the very features which make those things frightening or disturbing is exactly what adds to the value of those objects.

Moreover, by introducing the notion of “worth” when we ask whether it would be *worth* having a certain response to something, we must be careful that we are not illicitly moving between senses of fit, once again. One way of posing the question can’t help but stress the relevant prudential and moral considerations, for example, such as when someone claims that though a particular person’s obnoxiousness would be fitting to get angry at, that it isn’t *worth* getting angry at; or when someone claims that though certain video games might be entertaining, they are not *worth* one’s attention or engagement.

A second response might return to the thought that the relevant sense of fit here is not what I have been calling the fit of veridicality, but instead has something to do with *rationality*, understood in a particular way. While of course we can *understand* and *explain* why Jones Jr. has these negative attitudes toward reconciling with his father, this is not of central importance. The important question is whether he *justified* in having them, and it is commonly thought that in order to be justified, one must have a *justifying* reason – that is, something that speaks to the *goodness* of holding the attitude in question.

Note that Jones Jr. in fact seems well attuned and sensitive to the features of his circumstances. The nature of the demand that we might make of him to not be distressed or anxious about this conversation seems more like an expression that might help calm his nerves, rather than anything approaching a *rational* injunction or a claim about how well or badly he is thinking. Moreover, even if sense could be made of this claim as a rational injunction, we’d have reasons to be skeptical of the importance of the conception of rationality it presupposes. First,

this conception would seem to demand of us that our attitudes match our judgments of something's goodness or badness in a way that encourages *insensitivity* to the other properties of the object in question. Second, we'd have to claim that many rather mundane instances of a person's reaction to something – though out of step with her evaluative judgments – are instances of her irrationality. Again, laughing at a joke you know is not funny, or loving a film that you know is bad, for example, would count as a rational failing. Both of these commitments, I think, render this conception of rationality to be of limited importance.

I want to address one assumption of this last objection head on: the thought that the question worth pursuing here is whether a person is justified in reacting to the world in the way he does, in a sense of “justification” that is clearly distinct from, and more normatively significant than, a sense of “explanation”. As I’ve suggested, I think that the responses I’ve discussed – of Jones Jr., for example – *are* justified, in the sense that they are appropriately sensitive to the features of his situation: his anxiety, for example, *is veridical*. But even if there were an important sense of justification that I am giving short shrift to, according to which he (and others whose responses diverge from their evaluative judgments given the other features of their situation) should be considered unjustified or less than fully justified, it's nonetheless important to recognize the sense of fit in which our emotional responses need not be responsive to considerations of goodness and badness, but may still be appropriately held.¹⁴⁸ This is because we need the resources to distinguish the variety of distinct ways in which a person's responses can be fitting, or ill-fitting. We need to mark the difference between my nephew laughing at a

¹⁴⁸ The question of justification leads us back to the ambiguity that I first discussed. Justified, in what sense? *Rational* justification, and even talk of *reasons*, can seem a bit misplaced when it comes to some of the cases that I've discussed. Is it *rational* to be sad about a painful but necessary divorce with sadness? Would it be *irrational* not to be sad? It seems to involve a category error to even ask, let alone answer, these sorts of question. The awkwardness in assessing many of these cases in terms of either justificatory or (mere) explanatory reasons suggests to me that this dichotomy is one that we should resist, or at least amend: it isn't as helpful or illuminating as many have claimed.

fart joke, and him laughing at every acute angle he sees. And we need to mark the difference between someone's attitudes when they are intelligible but non-veridical from cases in which his attitudes are veridical. Even if Jones Jr. is criticizable (in some, not yet fully spelled out sense) for having negative attitudes toward something that he believes is good, we still need the resources to distinguish him from somebody who is made anxious by, for example, his teenager being a few minutes late for curfew or his spouse's not answering her cellphone, where in these latter cases there may be *nothing* that would make one's anxiety veridical, but the response is nonetheless intelligible. The importance of these features of our evaluative lives is overlooked and minimized by positions that assume that too sharp a distinction between the justified and the explained, and relegate all of these cases to the merely explained.

III. THE OTHER DIRECTION

I have so far argued against the claims that when it is fitting (in the sense of veridical) to have a positive or negative attitude toward x , x is good or bad, respectively. Many of our positive and negative attitudes, when fitting in the sense settled on, are fitting because of features of their objects that should not be understood as identical to or exactly supervenient on the goodness and badness of those objects, nor are they necessarily relevant when determining whether those objects are good or bad. And in Chapter 3, I argued that we can understand desire such that we can make sense of why there are things that are intelligible for a person to be attracted to, even when those things are bad – even morally bad – but because of other features of those things. Moreover, these things may not simply be intelligible for a person to desire, but even *veridical* to desire.

I will now focus on the other direction of the bi-conditionals stated at the beginning of the

paper. In particular, I will focus on the claim that if x is good then it is fitting to have a positive attitude toward x , leaving aside the parallel claim about badness. I will turn my attention away from discussing what it is that our desires or other positive attitudes might be responsive to, to questions about how it may be fitting to react to objects of goodness.

In particular, I will discuss three cases in which it seems possible to entertain the thought that something is good, while also not knowing how it would be fitting for a person to react to such a thing, or while also knowing that it *would* be fitting to have a negative attitude toward it, in regards to the very same feature of the object that makes it something good. Perhaps complicating this discussion is that I don't take myself to show, decisively, that the things under discussion here really *are* good. But the important question here is about the coherence and sensibility of maintaining, or entertaining such claims in the first place, and the philosophical questions that are fully intelligible once we see that certain claims about the goodness of something can be settled without that settling the question of how to respond to such things.

The Shape of One's Life

In *The Possibility of Practical Reason*, David Velleman objects to a standard conception of wellbeing according to which it is just the quantity of wellbeing accrued over a person's lifetime that determines the goodness of that life. He notes, for example, that we are more likely to think that a life in which one transitions from rags to riches is a better life than one in which one transitions from riches to rags, even if the amount of wellbeing in both lives is the same. And this, he suggests, is because of the importance of the *narrative significance* of one's life.

Let's entertain the thought that it is good for a person's life to have a good narrative arc. Now, imagine, for example, a person who is deliberating about a significant practical decision

that will effect how his life goes: should he propose to his childhood sweetheart, or not? Imagine that he takes into consideration things such as their compatibility, what it is that she wants out of life, and whether he is ready to settle down. Now, while mulling this over with a friend, imagine his friend states: *Well, it would make for a great story!*

While we can imagine the first man accepting this claim as well as the claim that it is good for a person's life to have a good narrative arc, we might think it strange for this fact to weigh with him, at all. It should not inform his deliberation, and it should not be something that makes the prospect of marrying his childhood sweetheart more attractive. One's own good narrative arc, though good, seems to not be the sort of thing that gives one reason to feel or do much of anything; or, if that claim is too strong: it would not only make sense, but may even be sensible, for one to recognize that one's narrative arc is good while at the same time having no positive attitudes toward it, whatsoever.

Hardship

It seems reasonable to think that hardship is good for human beings. We might think of how children who don't experience any hardship may grow up spoiled and unappreciative, or of how enduring struggle can develop one's character. Writer Alice Walker proposes the idea that it is *only* under certain forms of hardship – in particular, those that involve social ostracization and isolation – that a person can develop a valuable epistemic standpoint that is partly constituted by having experienced a life of hardship. She writes of her experience as a bullied child, “ugly” and “blind in one eye”:

...it was from this period – from my solitary, lonely position, the position of an outcast – that I began really to see people and things, really to notice relationships... the gift of loneliness is

sometimes a radical vision of society or one's people that has not previously been taken into account.¹⁴⁹

Nonetheless, given what hardship is actually like – it's *hard*, after all – we might think that it is fitting that a person would seek to avoid it, have a host of negative attitudes toward it, feel depressed or unhappy about it, and – except under controlled circumstances – be averse to seeking it out for those whom she loves. While it might be somewhat comforting for someone who is experiencing the hardship of poverty or alienation, for example, to be reminded that such hardship can be good, it wouldn't be unreasonable for her to find that this provides no emotional solace whatsoever. Alice Walker herself, reflecting on her difficult childhood, might be justified in feeling sorrow or bitterness, even though she recognizes that it was good for her to undergo it. And importantly, there seems to be no straightforward way to separate off the goodness of hardship from the features of it that make it something appropriately feared or avoided: it is the very fact that hardship is difficult and painful that makes it something valuable *and* something that it would be fitting to be averse to. Of course, this makes it a hard question to know to what extent hardship is good for a human being: while we see it as perhaps necessary for a good human life that a person undergo some hardship, we feel averse to letting our children and loved ones suffer from it at the same time. But this is a question that is left open for philosophical investigation should we see that settling that it is good is not enough to settle what to do about it.

Death

The final case I want to consider is from Bernard Williams' discussion of death in his article, "The Makropoulos Case: Reflections on the Tedium of Immortality". There, Williams argues that death is something that gives meaning to life. I won't assess the argument that he gives for this conclusion. Instead, I want to remark on some of the surrounding comments he

¹⁴⁹ *In Search of Our Mother's Gardens*, p. 264.

makes. He writes:

...death gives the meaning to life. That does not mean that we should not fear death (whatever force that injunction might be taken to have, anyway). Indeed, there are several very different ways in which it could be true at once that death gave the meaning to life and that death was, other things being equal, something to be feared... Considering whether death can reasonably be regarded as an evil is in fact as near as I shall get to considering whether it should be feared: they are not quite the same question.¹⁵⁰

Notice two of Williams' claims here. First, he notes that it could be possible that death is both something that gives meaning to life (that is, that it is something of quite deep value) and that – other things being equal – it is something to be feared. Second, he suggests that even if we were to conclude that death is an evil, this would still not settle the question of whether it should be feared. He is at pains to point out that these are different questions. Unsurprisingly, I think that Williams is right to distinguish them: questions about whether something is valuable should be separable from questions about how it would be reasonable to feel about it. Again, my suggestion is that this is because the features that make something a fitting object for certain attitudes need not be properties that are related to the *value* of those things. And lest we think that Williams' claims are intelligible only because he is focusing on different aspects of death – some that make it bad, and some that make it good – this need not be the case: the very thing that makes death valuable and important according to Williams (the annihilation of one's self) is the very thing that may make it something worth fearing.

Is it because we must think of death, or the cessation of the self, as both good *and* bad to generate this reaction? Even this does not have to be the case. Think for example, of the experience, referenced in countless films and books, of a person who is ready to die – who may even, because of a long illness, or as a noble sacrifice, or perhaps simply having lived a long,

¹⁵⁰ “The Makropulos Case: Reflections on the Tedium of Immortality”, p. 83.

good life, believes that it would be *good* for him to die – but who expresses fear or reservation about his own death. This fear, we might think, is fitting given what death is, even though his thinking his own death something good is also appropriate.

IV. CONCLUSION

When considering what we should conclude about a good narrative, death, and hardship, it seems to me that it would be fitting for a person to both entertain the thought that the object in question is something good *without* also having a positive attitude toward it. By this, not only do I mean that this would be an intelligible or permissible response; I also mean that it may be a sensible, or perceptive one. This is because, as I argued in the first part of this paper, many of our emotional and motivational responses are intelligent, but not value-centric. Furthermore, it seems to involve a sort of insight or acumen should a person recognize that hardship, for example, is something that can be good but nonetheless worth dreading, or that his dread, should he feel it, is no evidence that this hardship is genuinely bad in some respect, after all.

Perhaps at this point, a defender might grant that the claims cannot be as general as they are typically presented. This in itself would be something worth recognizing, given how often such claims are taken to be axiomatic! But it is worth considering how they might be altered to be made more plausible.

There are two different ways that we might interpret them, such that they are not subject to the worries that I have raised. The first is to limit the relevant positive and negative attitudes to a particular range of attitudes, more plausibly value-centric than those that I have discussed. The second is to limit the sense in which an object is said to have value, given that it would be fitting to have a positive or negative attitude toward it. We could formulate both claims in this fashion:

x is good iff x is a fitting object of a *limited* range of positive or pro-attitudes

(x is bad iff x is a fitting object of a *limited* range of negative or con-attitudes)

x is good *in some respect* iff x is a fitting object of a positive or con-attitude

(x is bad *in some respect* iff x is a fitting object of a negative or con-attitude)

A retreat to these interpretations, however, leaves us with claims too thin to be of much philosophical interest, if any. Let's consider the first. As we've seen, if this set includes the many attitudes that are typically thought of as positive or negative, this claim cannot be true. But which range of attitudes could we limit such claims to, such that the claim is vindicated? At first glance, we might think that some attitudes like *admiration* or *contempt* have this value-centric dimension. Then again, we might wonder whether it really is impossible to have contempt for somebody or something that one doesn't think of as bad, in any respect other than the thin, and already posited sense that one finds it fitting to have contempt for him. In any case, our investigation would best proceed in a piecemeal fashion – focusing on particular attitudes, and their particular veridicality conditions – rather than being granted by this general conceptual claim. But perhaps the more important point is this: Whatever the attitudes that populate this limited set, they will be so conceptually tied to the thought that the object is good that it becomes difficult to see how this claim could be asserting anything substantive about the relationship between value and our emotional and motivational lives.

Let's consider the second interpretation. In what respect is a thing good, given that it is a fitting object of a positive attitude? Of course, we should not conclude that it is *all things considered* good, for example. Perhaps it is just good to some limited extent. So, for example, we might judge it fitting for a scientist to find the ebola virus fascinating. So, we might conclude – not that the ebola virus is *good* – but that it is good to the extent that it is a good object of scientific interest. But again, this claim is too thin to be of much philosophical interest: it

amounts to no more than the very same point we've already made: that it would be fitting for a scientist to find ebola interesting.

Here is the dilemma for defenders of such claims: either such claims are to be read as including substantive judgments of value, or substantive positive and negative attitudes – in which case, as I've argued, they are sufficiently ambitious, but false: to think that something is good is not necessarily to be under any perceptual or normative or rational pressure to feel a substantive positive attitude toward it; nor does it imply the *unfittingness* of a variety of substantive negative attitudes. Or such views must retreat to versions of such claims that – while true – are trivially so.

The points I have made here highlight that there are at least three dimensions of third-personal evaluations that are overlooked or distorted when we render our emotions and motivations more value-centric than they really are. First, as I've already claimed, we are sensitive to when our responses are intelligible, even though they may not be veridical. Second, we are sensitive to the difference between a person's response being not only intelligible, but veridical – and we recognize that a response is often veridical, without being made so by the value or disvalue of its objects. And finally: there may even be virtues constituted by a person's disposition to experience her emotions in a non-value-centric way. Not only will we be able to arrive at more accurate understandings of our emotions by rejecting the two bi-conditionals I began with, we'll arrive at more accurate understandings of the evaluations that we make of one another in light of those emotions and motivations.

Furthermore, by refusing to limit our conceptions of value to our understandings of when it is and isn't appropriate for us to feel positively and negatively about things, we are able to expand our imaginations when it comes to investigating possible objects of goodness and

badness. This allows for us to consider the possibility of goods and bads, or evils, that we may otherwise lose sight of, and which – if we look exclusively to the objects of our appropriately held emotions and motivations – we may never fully recognize or attend to. For example, we should be able to question whether or not a person’s death is good or bad, even after we’ve settled how to feel about it.

Finally, by rejecting the dogma I’ve focused on here, we are able to take into more serious consideration a certain understanding of value that allows for goodness and badness to exist, at least in principle, independently of our positive or negative feelings toward it, even when those feelings are appropriately held. Take, for example, Robert Adam’s discussion of transcendent goodness. Adams writes:

The idea of the Holy... is an idea of something that is wonderful in a way that escapes both human understanding and human purposes... the Holy is love... but ‘nice’ is definitely not the word for it. From a human point of view, the Holy has rough edges. It screams with the hawk and laughs with the hyenas. We cannot comprehend it. It is fearful to us, and in some ways dangerous. All of this, on the theistic view that I favor, is true of the Good itself. We are not the measure of all things, and have at best a very imperfect appreciation of the full dimensions of the good.¹⁵¹

Like the previous examples I discussed above, Adams’ discussion of goodness depends upon an alternative way of thinking about what goodness could be, one that is ruled out by the claims that I have here argued against. This alternative allows the Good is something that it may be fitting for us to have certain negative attitudes toward, in spite of our recognition that those things are good: goodness really is transcendent of us, and of even our justifiably felt attitudes. And it is an alternative that is worth exploring, and may be particularly well suited to capture the ways in which we experience the good and the bad while also leaving room for more accurate understandings of human emotion and motivation than more limited frameworks can provide. It

¹⁵¹ *Finite and Infinite Goods*, p. 52.

should, at any rate, not be something ruled out as conceptually confused or incoherent on the basis of two abstract claims that – though funding many interesting and ambitious theoretical projects – do so at the cost of distorting the nature of our emotions, and our relationship to goodness.

Conclusion

I. INTRODUCTION

My dissertation has responded to a picture of human psychology according to which our judgments or beliefs about value on the one hand, and our motivations and emotions on the other, are best made harmonious with one another. According to the simplest version of this picture – one that I described as a descriptive, rather than a normative or ethical claim – to judge or believe that something is good just *is* to respond to it with a positive response such as attraction, desire, love, or promotion. This thought is coupled with a parallel claim about judgments that something is *bad*, and negative responses like repulsion, aversion, or hatred.

As I have illustrated throughout, this basic position and its variants are widespread and maintained by what are otherwise vastly different philosophical views. For example, though a particularly robust unification is commonly associated with a champion of the intellect, Socrates, this unity is also a core commitment of views that deny that sincerely held evaluative beliefs or judgments can exist independently of a person's desires and emotions. Hare's prescriptivism commits to a form of unification by insisting that no sincere avowal of an evaluative judgment can be made without some congruent desire, or emotion, since such a claim amounts to stating an

imperative.¹⁵² And forms of expressivism or quasi-realism, too, must maintain that there is a close connection between our evaluative judgments and our feelings.¹⁵³

The motivations for these two latter positions are meta-ethical, having to do with the metaphysical and epistemological consequences of treating evaluative knowledge just like knowledge in other domains. But one might find such a unification compelling for more familiar, everyday reasons. While there are numerous interpretations of why it is that Socrates defended his intellectualism, one reason was ethical, or moral in nature: he wanted to emphasize the *importance* of evaluative knowledge in living a good life. For Socrates, possessing genuine knowledge about the value of things is not only necessary, but is also *sufficient* for, doing the right things and feeling the right sorts of ways. This old idea has re-emerged in contemporary philosophical literature, maintained by those who defend the thought that moral virtue *just is* moral knowledge, of a sort. To truly know the good is to love it; to act badly or wrongly is to demonstrate that one's grasp of the good is only partial, hazy, superficial, or distorted.¹⁵⁴

While there may be a very particular sense in which knowledge of the good guarantees this emotional or motivational response – a sense that would capture at least some of our thoughts about why some people fail to act or respond in the ways we'd hope – there are nonetheless familiar ways in which a person can be fully aware of what is good or bad, while experiencing less harmonious emotions or motivations in response. Not only is this sort of response possible, I've argued that depending on the particularities of that disharmony, it may even reflect something good about the person who experiences it. So I, too, have hoped to discuss the intersection at which these psychological claims overlap and interact with moral, and

¹⁵² R. M. Hare, *Freedom and Reason*.

¹⁵³ Simon Blackburn, *Ruling Passions: A Theory of Practical Reasoning*.

¹⁵⁴ For example, see John McDowell, "Are Moral Requirements Hypothetical Imperatives?", Martha Nussbaum, "The Discernment of Perception: an Aristotelian Conception of Private and Public Rationality," in *Love's Knowledge: Essays in Philosophy and Literature*, and David McNaughton, *Moral Vision*.

more broadly ethical, questions. I will now conclude this investigation by elaborating on some of the conclusions that I have hoped to argue toward, making explicit their connections to other positions and approaches in moral psychology and ethics.

II. AN ADEQUATELY COMPLEX PSYCHOLOGY

In chapter 3, I challenged the unification thesis by arguing against one of its most minimal, and so most plausible, variations – the guise of the good thesis. But even showing that this claim is false leaves open another which I think should also be rejected: one according to which it is possible to knowingly desire or love the bad and hate or be uninterested in the good, but that this occurs, necessarily, as a failure in *rational agency*.

There are, of course, various ways to understand what being a “rational agent” amounts to. Sometimes this claim results simply from overstating the importance of psychic coherence in and of itself as an essential aspect of being rational.¹⁵⁵ But this thought is also bolstered by substantive psychological views with substantive understandings of rationality.

These views characterize such disunities as best understood as, for example, (i) necessarily a matter of one’s reason being overcome or overwhelmed by a passion, (ii) necessarily a matter of one’s emotions or motivations undermining or distorting one’s understanding or perception at the time of action, or (iii) necessarily resulting in psychological states and actions that cannot be understood as one’s own.

¹⁵⁵ There are probably particular forms of coherence that *are* important for rationality, but it is often treated as a self-evident, blanket truth that the more all of a person’s mental states cohere with one another, the more rational she is. Even Nomy Arpaly, who has argued for the rationality of acting against one’s better judgment, nonetheless also accepts that in general, the more coherent your desires are with one another, the more rational you are. But this claim needs to be spelled out more precisely, if it is to be plausible: it doesn’t strike me as necessarily irrational, for example, to both want and not want something, even for the very same reason; nor does it strike me as irrational to find oneself wanting something very badly in spite of its conflicting with the other desires that one has – such as wanting very badly to quit one’s job and start over, even though one would also like to succeed at one’s job. If coherence is important, it will be particular forms of coherence that will need to be distinguished from other, less important forms.

As an example of the first sort of view (according to which disunities are best understood as a matter of one's reason being overcome or overwhelmed by passion), take Gary Watson's preferred picture of the soul, which he takes to be roughly Platonic. Watson writes,

Plato was calling attention to the fact that it is one thing to think a state of affairs good, worth while, or worthy of promotion, and another simply to desire or want the state of affairs to obtain. Since the notion of value is tied to (cannot be understood independently of) those of the good and worthy, it is one thing to value (think good) a state of affairs, and another to desire that it obtain. However, to think a thing good is at the same thing to desire it (or its promotion). Reason is thus an original spring of action. It is because valuing is essentially related to thinking or judging good that it is appropriate to speak of the wants that are (or perhaps arise from) evaluations as belonging to, or originating in, the rational (that is, *judging*) part of the soul; values provide *reasons* for action. The contrast is with desires, whose objects may not be thought good and which are thus, in a natural sense, blind or irrational.¹⁵⁶

Though Watson suggests that it is impossible to think a thing good without desiring it or its promotion, he does allow for disunity in the other direction: it is possible to desire something without thinking it good. To illustrate, he provides cases in which (he proposes) a person desires something that she does not believe good – “a woman who has a sudden urge to drown her baby in the bath; the case of a squash player, who, while suffering an ignominious defeat, desires to smash his opponent in the face with a racquet.”¹⁵⁷ For Watson, pure or simple desires – uninformed by Reason and so uninformed by considerations of value – are “blind” or “irrational”, and as his selection of examples suggest, we can be *overcome* by such desires.¹⁵⁸

As an example of second kind of picture (according to which akratic disunities are a matter of one's emotions or motivations undermining or distorting one's understanding or

¹⁵⁶ "Free Agency," p. 208.

¹⁵⁷ *ibid.* 210.

¹⁵⁸ There are actually two separate issues here that one might think are essential to making a desire “rational”: one is to think that it must be sensitive to value, another is to think that it must be sensitive to the agent's own judgments, or deliberation. I did not discuss this second claim, but I don't think that it is true – an agent can act rationally, or intelligently, without deliberating.

perception at the time of action), take Richard Kraut's discussion of how Aristotle understands the possibility of *akrasia*. According to Kraut's reading, the akratic agent, given that her desires or emotions motivate her against what she knows would be best, does not *fully* grasp what would be best to do at the time in which she is motivated to do what would not be best. Kraut writes,

Aristotle [compares] *akrasia* to the condition of other people who might be described as knowing in a way, but not in an unqualified way. His examples are people who are asleep, mad, or drunk; he also compares the akratic to a student who has just begun to learn a subject, or an actor on the stage (1147a10–24). All of these people, he says, can utter the very words used by those who have knowledge; but their talk does not prove that they really have knowledge, strictly speaking... The akratic says, at the time of action, that he ought not to indulge in this particular pleasure at this time. But does he know or even believe that he should refrain? Aristotle might be taken to reply: yes and no. He has some degree of recognition that he must not do this now, but not full recognition. *His feeling, even if it is weak, has to some degree prevented him from completely grasping or affirming the point that he should not do this...* When reason remains unimpaired and unclouded, it dictates will carry us all the way to action, so long as we are able to act.¹⁵⁹

On this basic picture, the best metaphors for disunity in the soul are not those of mutiny or overthrow by one's passions, but instead of one's emotions and desires *clouding* or *distorting* reason so that while there is a sense in which we know what would be better to do, it is not enough to "carry us all the way to action". Particular attractions obscure or undermine our grasp of something's *true* value: in Watson's example, my anger at my squash opponent for calling another let clouds my knowledge or awareness of the fact that I ought not strike him with my racquet, even though I "know" I ought not. Such aspects of our psychology should be understood then, as necessarily adverse to our rationality, at least at the time of akratic action.

Finally, as an example of the third kind of picture, (according to which akratic disharmonies necessarily give rise to actions that cannot be understood as one's own, or as resulting from psychological states that cannot be understood as one's own) take Sarah Buss's

¹⁵⁹ Richard Kraut, "Aristotle's Ethics", *Stanford Encyclopedia of Philosophy*, April 2014. Jessica Moss has also defended this understanding of *akrasia*, both as a reading of Aristotle and Plato, and as itself a good understanding of the phenomenon.

neo-Kantian account of the constraints on genuinely intentional action. According to Buss, an agent's evaluative beliefs must play a direct role in an agent's action if we are to understand it as resulting from reasons that are "her own". According to Buss, intentional agents "must *endorse* their actions [as good or worth performing] *in addition to desiring* to perform them", where non-intentional actions, in being unconditioned by such evaluative beliefs, are in an important sense external to the agent.¹⁶⁰ If I don't endorse my attack as *good* in some way, it wasn't really *me* performing the action – my anger has compromised my ability to perform full-fledged actions.

I think that these alternatives, though allowing for certain kinds of disunity, nonetheless give us an incomplete picture. To be clear, I am not denying that each of these views describe genuine psychological phenomena. We can be overcome by certain parts of our psychology, when we are overcome we are sometimes irrational because of this; our desires and feelings can cloud our judgment of how valuable something is; and there are aspects of our psychologies that are best understood as not being truly our own – including, in some cases, the desires or attractions that we have toward things we simultaneously think of as worthless, or bad.

The worry, however, is that by adopting these views of human psychology, we too quickly conclude that having one's emotions and motivations diverge from one's evaluative judgments or beliefs is *necessarily* irrational or compromising of one's agency – and it is this conclusion that I've argued is mistaken. It is a mistake that I think can be explained by one widespread assumption that our emotions and desires are what I called *value-centric* – appropriately or rationally held only when directed toward apparent goodness or badness.¹⁶¹ We

¹⁶⁰ Buss, "What Practical Reason Must Be If We Act For Our Own Reasons", p. 412. Buss writes, "The claim is... that if *qua* deliberator, the agent is against performing her action [that is, if she believes that what she is doing is not good to do] then, *qua* deliberator, i.e., *qua* initiator of her own action, *she* has no motivation to perform it."

¹⁶¹ As Jessica Moss has put this point in her reading of Plato, "All passions (desires, emotions, pleasures and pains) are responses to things *qua* valuable, but only calculation can grasp what is truly good or bad, as opposed to what merely appears so," "Appearances and Calculations: Plato's Division of the Soul", p. 7.

have a tendency to think that our motivations and emotions must *either* be value-centric, or they must be “blind” as Watson suggests. In contrast to both of these conclusions, it seems that our desires and many of our emotions can be understood as responses to features of an agent’s circumstances that should not be understood as identical to features of value or disvalue. And when we assess those desires or emotions as “fitting” or not, we are sometimes only interested in whether they are sensitive to *those* features – it’s a further question whether or not they are sensitive to what is good, or bad. For example, as I discussed in Chapter 4, being *interested* in something is a paradigmatically positive response, but is not exclusively sensitive to considerations of positive value. One’s interest then, can be justified or veridical, as long as the object that sustains interest is, in fact, interesting. We might then ask: would it be *appropriate* for one to be interested by this object? For example, is it *good* to be interested in the car crash down the street? This sort of question presupposes the thought that one’s interest might *not* be appropriate, given the badness of the event. But importantly, this thought brings with it a different sense of “fit”, or “appropriateness”, that is distinct from the first.

Once we see that there are many emotions that this alternative sort of justification or fit is apt, we begin to lose the sense that something has necessarily gone wrong at the level of clear thought and perception when they are not congruent with our judgments or experiences of *value*. The rubbernecker might be *morbid* or *inconsiderate*, but not “irrational”. Moreover, this allows us to see that such attitudes may be recalcitrant to considerations of value, without this being evidence of irrationality – they *appear* to be recalcitrant because they are not exclusively sensitive to those considerations to begin with. To ask Glory, for example, to reflect again on the *badness* of her brother’s character, or to present to her with more evidence of its badness, may not change her attitude toward him – not because her love is irrational or arational – but because

it is not exclusively informed by considerations of goodness and badness of its object, but something else.

Given these considerations, we should also lose at least some of the motivation we have for immediately disowning, or externalizing, these attitudes as not properly part of a person's self. Part of what makes these attitudes easy to disown is that they seem so difficult to square with what we think makes us rational, intelligent, and responsive agents. But when we see that our conclusions have resulted from a misunderstanding of their nature, they become more familiar as properly internal aspects of our psychology even when their objects are not good.

III. PLURAL AND CONFLICTING VIRTUES

The psychological claims I've defended are reinforced and more fully illustrated once we move to the more explicitly ethical questions that arise when considering cases of psychological division. I've argued that depending on the way in which a person's psychology is divided, such a division may constitute or be otherwise inseparable from virtues that we are more likely to overlook by venerating the unified mind.¹⁶²

I've challenged the idea that harmony is *essential* to being a good person – arguing that there are a number of virtues which are partly constituted or inseparable from a person's motivations and emotions diverging from her evaluative judgments – depending on the particular way in which that division manifests itself. These virtues include a form of graciousness, self-

¹⁶² It might now be easier to discuss how my conclusion differs from the one argued for by Nomy Arpaly in, for example, "On Acting Rationally Against One's Better Judgment". Our positions differ in at least three different respects. First, Arpaly is interested in assessing particular *actions* as either rational or irrational for a person to perform, whereas I am interested in defending a disposition to act against one's better judgment. Second, her defense depends on the action's being *justified* – though not to the agent's knowledge at the time of the action, whereas I am arguing for the value of the disposition, even though it gives rise to particular actions that may not be justified according to the standard that she uses. Third, Arpaly accepts a broadly Humean picture of the mind, and action – one that I've opposed here.

esteem, humility, solidarity with one's fellow human beings, vivacity, perspicacity, and particular forms of intelligence.

Do I mean to *deny* that it is a mark of a good life that a person's psychology is unified? In one sense, yes. Unification on its own may not be a mark of a good life if such unification comes at a cost of being adequately sensitive to one's circumstances. Such unification harmonizes the interior life of an agent at the cost of her engagement with and perception of the world. It might also be partly explained by a person's *lack* of the virtues I've described – a particularly arrogant person, for example, can maintain a constancy between his desires and his evaluative beliefs by simply taking the latter to be settled by the former.

But of course, *that* sort of harmony is unlikely to be what Stocker, and others, meant to be highlighting as so important to living well. So we should recognize that it may be psychologically possible to remain unified even in spite of one's full sensitivity to the features of one's situation that would justify division instead, such as when the best thing for a person to do is also something exceedingly boring, or painful, or annoying to do. Rather than being an *effort* to always love the good and hate the bad, this sort of person's responses might come quite easily to her. In fact, this is what it might take to be the truly virtuous agent on a broadly Aristotelian account.

Such people – depending on the particular circumstances – might be exceptionally stoic, patient, exacting, or serious. I do not deny that should such psychologies be possible for human beings to possess, that these beings would possess certain important virtues. Nor do I deny that the people that I've picked out as examples of possessing the virtues of disunity might possess certain vices, or character traits that it would be reasonable to be ambivalent about. We might think that Sherlock Holmes' *intensity* is a virtue of his, even if it comes at the cost of his being,

for example, more patient or sanguine. And Emma Woodhouse's playfulness in conversation is something that leads easily to hurt feelings and misunderstandings.¹⁶³

But the point that I have tried to make here is that even so, given what the virtues that I've described are like, there may be no fact of the matter which sort of psychological state – harmonious or divided – is simply all things considered *better* than the other. Instead, the observations I've made suggest moving toward a model of the virtues that does not fit well with this tendency to rank them against one another – but instead appreciates them each as different, but nonetheless excellent, characteristics for a person to possess. There may simply be no answer to the question of whether Sherlock or Watson is, all things considered, the better person overall. This seems to me, to be a good result – one that rightly celebrates a diversity of human excellences.

Now, one might wonder why it is that these virtues have been overlooked or overshadowed by those that more obviously depend on a unified psychology. One suggestion is not that we have privileged harmony too much, but instead, that we haven't fully recognized the implications of a thoroughly pluralistic account of *value*. We might, for example, want to maintain the basic thought that genuine virtues are necessarily partly constituted by a person's having a unified psychology, but pair this with the observation that there are more good things than philosophers are typically willing to recognize.

While this conclusion would preserve the value of a unified psychology, the worry I raised is that by endorsing this alternative picture we risk distorting our substantive evaluations. There, I suggested that we have good reason for maintaining the thought that a human being's "streak of malice" or "unaccountable vanity" *are* genuinely bad aspects of his character, even though there is something to be said for not solely lamenting or regretting these features of who

¹⁶³ "Badly done, indeed!"

he is. Furthermore, we also have good reason to keep distinct genuine excellences *as* excellences – ones that are importantly different from aspects of human nature that, though able to inspire affection, may not be all that good. That there is something good about people who knowingly love sentimental kitsch gives us little reason to elevate the evaluative status of velvet portraits of Elvis or porcelain statuettes of saucer-eyed puppies.

Moreover, we might again fall into the mistake of treating our positive and negative emotional responses as more value-centric than they really are – and this would encourage a distortion of those attitudes. And finally, as I discussed in my final chapter, I think that allowing for a more flexible understanding of the appropriate responses to things both good and bad, rather than a more flexible understanding of what sorts of things are good or bad, will better accommodate the phenomenology of at least some experiences of both.

So although I think we ought to be flexible when considering what things are of genuine value, this sort of pluralism must also be supplemented by another: a pluralism of virtue. This is not just a denial of one way of understanding the unity of the virtues – it is not, for example, to simply point out the possibility that a person might be brave without being generous, or kind without being prudent. Rather, the thought is that not only is it possible to possess certain virtues without possessing the others, there may be some that genuinely *conflict* with one another. It may not be possible to possess all of the virtues that are partly constituted by loving the good, while possessing those that are partly constituted by loving the bad. There may not be just *one*, maximally good or excellent way to be. And again, this seems to me to be a good result.

IV. VIRTUES & THE IMPORTANCE OF EVALUATIVE KNOWLEDGE

At this point, one might wonder whether I've established the result that there are virtues that are partially constituted by psychic disharmony by antecedently committing to a picture of

the virtues as good or attractive but relatively minimal qualities of a person's character.¹⁶⁴ For example, according to Hume, virtues are simply those mental qualities that are either immediately agreeable to or useful for their possessors, or for others. Given some of the examples I've used, it may seem that is precisely the sort of account I've assumed. My discussion of David Chang, for example, was predicated on the observation that there is – at least for some of us – something immediately appealing, or attractive, about the way in which his desires are not aligned with his evaluative knowledge.

Or take, for example, Julia Driver's consequentialist conception of the virtues, according to which a virtue is simply, "a character trait ... which, generally speaking, produces good consequences for others"; a vice is simply a trait that, "generally speaking ... leads to bad effects."¹⁶⁵ When compared to a standard Aristotelian understanding of the virtues, these alternative understandings de-emphasize the internal life of the virtuous agent. They focus our attention instead on the *effects* of such states. Because of this, they accommodate a number of traits that more stringent accounts wouldn't *count* as genuine virtues, or vices. This "thinness" is precisely why some philosophers, such as Rosalind Hursthouse¹⁶⁶ and Robert Adams¹⁶⁷, have been adamant in maintaining that genuine virtues must be psychologically deeper, and more complex, than the qualities that Hume and Driver would allow.

I think that both Hume and Driver are right for highlighting the psychologically thinner dispositions that we enjoy and love in one another – and which are likely to be much more important in our lives than many philosophers may be willing to grant. We *do* rightly appreciate

¹⁶⁴ One might also be waiting for me to provide a complete account what it is for a quality to be a virtue – but I won't be doing so here.

¹⁶⁵ Julia Driver, *Uneasy Virtue*, pp. 123–124.

¹⁶⁶ Rosalind Hursthouse, "Are Virtues the Proper Starting Point for Morality?".

¹⁶⁷ Robert M. Adams, *A Theory of Virtue: Excellence in Being for the Good*.

that a person has perfect pitch, or the ability to load the dishwasher perfectly. Nonetheless, I want to finish my discussion here by emphasizing how my conclusions are not, in fact, all that friendly to these more minimalistic accounts of virtue. One rather small point is that I am skeptical of the ability of either of these alternative accounts to capture at least some of the virtues that I am interested in, given their focus on qualities that are appealing, or pleasant, or useful, or generally have good consequences. Some of the features I've picked out do not strike me as necessarily benefitting or being useful for others or for their possessors. Moreover, one would be justified in finding at least some of these qualities quite *annoying*, *frustrating*, or *abrasive* – though nonetheless good.

But suppose that these views can accommodate such virtues, after all. Nonetheless, there does seem to be an important difference between the positive qualities of character that Hume and Driver would insist are genuine virtues, and the kinds of quality of character that Aristotle and others are interested in picking out. These latter qualities are psychologically more demanding, and because of this, reveal to us more about the agent's inner life as well as her experience of the world. I think that this is an important quality to preserve as distinct and distinctively important. So while a disagreement about what is essential to virtue might strike some as an uninteresting terminological dispute, it's important to notice that there are substantive differences between the kind of qualities that Hursthouse and Adams are interested in, and the kind of qualities that Hume and Driver are willing to count as virtues.

And the qualities that I hope to have picked out do, I think, have this depth to them. In particular, it was important for my conclusions that the people I've discussed *do* have the relevant sort of evaluative knowledge and perception. This is what marks them out as different from those who possess similarly valenced desires and emotions, but do so either in evaluative

ignorance, evaluative arrogance, or without consideration about the value of the relevant things. For example, as I discussed in Chapter 4, a person's susceptibility to boredom is something we appreciate more when we are convinced that her boredom is accompanied by her awareness of the *value* of the things around her – otherwise we may suspect she is being inattentive, or shallow, or ungrateful. Part of one's appreciation of Chang's response is that we know that he knows what sorts of food are excellent and why, but that he just doesn't *care* all that much about eating only excellent food. And the sort of love I discussed in Chapter 2 is, like with other excellent forms of love, a love that isn't blind to the bad qualities in its objects: it is one that is exercised with discernment and sensitivity. So in all of the cases I've discussed, there is something that would be lost should we fail to see how it is that the person sees the world, and her emotions and motivations as expressions of her particular stance within it.

And so, although I began in stark disagreement with one old Socratic claim, I've reinforced another: the importance of evaluative knowledge. Importantly, we shouldn't take this to mean that such knowledge is sufficient to guarantee a harmonious psychology, or that it will be clear what to do or how to feel once we have it, nor should we think that this lack of guarantee is something regrettable. It is within this diversity of responses to the good and the bad that we will find and recognize particular patterns of response worth exploring – and perhaps more importantly, celebrating.

REFERENCES

- Adams, Robert, M. 1972. "Must God Create the Best?" *Philosophical Review* 81 (3).
- Adams, Robert, M. 2002. *Finite and Infinite Goods: A Framework for Ethics*. Oxford: Oxford University Press.
- Anderson, Elizabeth. 1993. *Value and Ethics in Economics*. Cambridge, Mass: Harvard University Press.
- Anderson, Elizabeth. 1997. "Practical Reason and Incommensurable Goods." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang. Cambridge, MA: Harvard University Press.
- Annas, Julia. 1993. *The Morality of Happiness*. Oxford: Oxford University Press.
- Anscombe, G.E.M. 1957. *Intention*. Ithaca, NY: Cornell University Press.
- Aristotle. 1999. *Nicomachean Ethics*. 2nd ed. Indianapolis: Hackett Publishing Co.
- Arpaly, Nomy. 2000. "On Acting Rationally Against One's Best Judgment." *Ethics* 110:488-513.
- Austen, Jane. 1816. *Emma*. London: John Murray.
- Bagley, Benjamin. 2015. "Loving Someone in Particular." *Ethics* 125 (2):477-507.
- Berryman, John. 1965. *77 Dream Songs*. New York: Farrar, Straus and Giroux.
- Blackburn, Simon. 1998. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press.
- Boyle, Matthew, and Douglas Lavin. 2010. "'Goodness and Desire'." In *Desire, Practical Reason, and the Good*, edited by Sergio Tenenbaum. New York: Oxford University Press.
- Bradford, Gwen. 2015. *Achievement*. Oxford: Oxford University Press.
- Brewer, Talbot. 2011. *The Retrieval of Ethics*. Oxford: Oxford University Press.
- Broome, John. 2007. "Does Rationality Consist in Responding Correctly to Reasons?" *Journal of Moral Philosophy* 4:349-74.
- Buss, Sarah. 1999. "What Practical Reasoning Must Be If We Act for Our Own Reasons." *Australasia Journal of Philosophy* 77:399-421.
- D'Arms, J, and D Jacobsen. 2003. "The significance of recalcitrant emotion (or, anti-quasijudgmentalism)." In *Philosophy and the Emotions*, edited by Anthony Hatzimoysis. Cambridge: Cambridge University Press.

- Davidson, Donald. 1963. "Actions, Reasons, and Causes." *The Journal of Philosophy* 60 (23):685-700.
- Davidson, Donald. 1970. "How is Weakness of the Will Possible." In *Moral Concepts*, edited by Joel Feinberg. London: Oxford University Press.
- Doyle, Arthur Conan. 1892. "A Scandal in Bohemia." In *The Adventures of Sherlock Holmes*. London: George Newnes.
- Doyle, Arthur Conan. 1908. "The Singular Experience of Mr. John Scott Eccles." *The Strand Magazine*.
- Driver, Julia. 2001. *Uneasy Virtue*. Cambridge: Cambridge University Press.
- Feinberg, Joel. 2003. *Problems and the Roots of Law*. Oxford: Oxford University Press.
- Fitz-Gerald, Sean. 2015. "61 Comedians Recall their Favourite, First, and Life-Changing Jokes." *The Vulture*.
- Foot, Philippa. 2002a. "Goodness and Choice." In *Virtue and Vices: And Other Essays in Moral Philosophy*. Oxford: Oxford University Press.
- Foot, Philippa. 2002b. "Virtues and Vices." In *Virtues and Vices: and Other Essays in Moral Philosophy*. Oxford: Oxford University Press.
- Frankfurt, H. 2002. *Contours of Agency: Essays of Themes from Harry Frankfurt*. Edited by Sarah Buss and Lee Overton. Cambridge, MA: MIT Press.
- Frankfurt, H. 2004. *The Reasons of Love*. Princeton: Princeton University Press.
- Frankfurt, Harry. 2006. *Taking Ourselves Seriously and Getting it Right*. Edited by Debra Satz. Stanford: Stanford University Press.
- Gauthier, David. 1986. *Morals by Agreement*. Oxford: Oxford University Press.
- Gewirth, Alan. 1980. *Reason and Morality*. Chicago: University of Chicago Press.
- Hardy, Thomas. 2001. *Tess of the D'Urbervilles*. New York: Dover.
- Hare, R. 1963. *Freedom and Reason*. Oxford: Oxford University Press.
- Hawkins, Jennifer. 2008. "Desiring the Bad Under the Guise of the Good." *The Philosophical Quarterly* 58 (231):244-264.

- Herman, Barbara. 1996. "Making Room for Character." In *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, edited by Stephen Engstrom and Jennifer Whiting. Cambridge: Cambridge University Press.
- Hurka, Thomas. 2003. *Virtue, Vice, and Value*. New York: Oxford University Press.
- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. New York: Oxford University Press.
- Hursthouse, Rosalind. 2006. "Are Virtues the Proper Starting Point for Morality?" In *Contemporary Debates in Moral Theory*, edited by James Dreier. Malden, MA: Blackwell Publishing.
- Kolodny, Niko. 2003. "Love as Valuing a Relationship." *Philosophical Review* 112 (2):135-189.
- Kontos, Pavlos. 2014. "Non-Virtuous Intellectual States in Aristotle's Ethics." In *Oxford Studies in Ancient Philosophy*, edited by Brad Inwood. Oxford: Oxford University Press.
- Korsgaard, Christine. 1986. "Kant's Formula of Humanity." *Kant-Studien* 77 (1-4).
- Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Kraut, Richard. 2014. Aristotle's Ethics. Stanford Encyclopedia of Ethics.
- MacFarquhar, Larissa. 2008. "Chef on the Edge." *The New Yorker*.
- MacFarquhar, Larissa. 2015. *Strangers Drowning: Grappling with Impossible Idealism, Drastic Choices, and the Overpowering Urge to Help*. New York: Penguin Press.
- McDowell, John. 1979. "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society, Supplementary Volumes* 52:13-29.
- McNaughton, David. 1988. *Moral Vision: An Introduction to Ethics*. Malden, MA: Blackwell Publishers.
- Moore, G.E. 2004. *Principia Ethica*. Mineola, New York: Dover Publications.
- Moss, Jessica. 2008. "Appearances and Calculations: Plato's Division of the Soul." In *Oxford Studies in Ancient Philosophy*.
- Murdoch, Iris. 1971. *The Sovereignty of Good*. New York: Schocken.
- Murphy, Jeffrie. 1982. "Forgiveness and Resentment." *Midwest Studies in Philosophy* 7.
- Nietzsche, Friedrich. 2003. *The Genealogy of Morality*. New York: Dover Publications.

- Nussbaum, Martha. 1990. "The Discernment of Perception: An Aristotelian Conception of Private and Public Rationality." In *Love's Knowledge: Essays on Philosophy and Literature*. New York: Oxford University Press.
- O'Brien, Wendell. 2014. "Boredom." *Analysis* 74 (2):236-244.
- Oddie, Graham. 2009. *Value, Reality, and Desire*. Oxford: Oxford University Press.
- Parker, Dorothy. 1976. "Observation." In *The Portable Dorothy Parker*. New York: Penguin Books.
- Pettigrove, Glen. 2012. *Forgiveness and Love*. Oxford: Oxford University Press.
- Prejean, Helen. 1993. *Dead Man Walking*. New York: Random House.
- Preston-Roedder, Ryan. 2013. "Faith in Humanity." *Philosophy and Phenomenological Research* 87 (3).
- Quinn, Warren. 1993. "Putting Rationality in its Place." In *Morality and Action*. Cambridge: Cambridge University Press.
- Rabinowicz, Wlodek, and Rønnow-Rasmussen. 2004. "The Strike of the Demon: On Fitting Pro-Attitudes and Value." *Ethics* 114:391-423.
- Railton, Peter. 1997. "On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action." In *Ethics and Practical Reason*, edited by G. Cullity and B. Gaut, 53-79. New York: Oxford University Press.
- Rawls, John. 1999. *A Theory of Justice*. Cambridge: Belknap Press.
- Raz, Joseph. 2010. "On The Guise of the Good." In *Desire, Practical Reason, and the Good*, edited by Sergio Tenenbaum. New York: Oxford University Press.
- Reath, Andrews. forthcoming. "Did Kant Hold that Rational Volition is Sub Ratione Boni?" In *Reason, Value and Respect: Kantian Themes from the Philosophy of Thomas E. Hill Jr.*, edited by Robert Johnson and Mark Timmons. Oxford: Oxford University Press.
- Robinson, Marilynne. 2004. *Gilead*. New York: Farrar, Straus and Giroux.
- Robinson, Marilynne. 2008. *Home*. New York: Farrar, Straus & Giroux.
- Ross, W.D. 1939. *Foundations of Ethics*. Oxford: Clarendon Press.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

- Scanlon, Thomas. 2011. "The Unity of the Normative." *Philosophical Studies* 154 (3):443-450.
- Schafer, Karl. 2013. "Perception and the Rational Force of Desire." *Journal of Philosophy* 110 (5):258-281.
- Setiya, Kieran. 2007. *Reasons without Rationalism*. Princeton: Princeton University Press.
- Sinhababu, Neil. 2009. "The Humean Theory of Motivation Reformulated and Defended." *The Philosophical Review* 118 (4):465-500.
- Smith, Michael. 1987. "The Humean Theory of Motivation." *Mind* 96 (381):36-61.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell
- Sontag, Susan. 1961. "Notes on Camp." In *Against Interpretation*. New York: Deli Laurel.
- Stampe, Dennis. 1987. "The Authority of Desire." *Philosophical Review* 96 (3):335-381.
- Stocker, Michael. 1976. "The Schizophrenia of Modern Ethical Theories." *The Journal of Philosophy* 73 (14):453-466.
- Stocker, Michael. 1979. "Desiring the Bad: An Essay in Moral Psychology." *The Journal of Philosophy* 76 (12):738-753.
- Stocker, Michael. 2004. "Raz on the intelligibility of bad acts." In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace. Oxford: Oxford University Press.
- Strawson, P.F. 1974. "Freedom and Resentment." In *Freedom and Resentment and Other essays*. London: Methuen.
- Sussman, David. 2009. "For Badness' Sake." *Journal of Philosophy* 106 (11).
- Tenenbaum, Sergio. 2007. *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge: Cambridge University Press.
- Tenenbaum, Sergio. 2010. *Desire, Practical Reason, and the Good*. Edited by Sergio Tenenbaum. Cambridge: Cambridge University Press.
- Thorpe, John. 2003. "Aristotle on Brutishness." *Dialogue* 42.
- Velleman, David. 1992. "The Guise of the Good." *Nous* 26 (1):3-26.
- Velleman, David. 1999. "Love as a Moral Emotion." *Ethics* 109 (2):338-74.
- Wallace, David Foster. 2011. *The Pale King*. New York: Little, Brown and Company.

- Wallace, R. Jay. 1999. "Three Conceptions of Rational Agency." *Ethical Theory and Moral Practice* 2.
- Waters, Alice. 2013. Alice Waters on Talk to Al Jazeera. edited by David Shuster: Al Jazeera America.
- Watson, Gary. 1975. "Free Agency." *The Journal of Philosophy* 72 (8).
- Wiggins, David. 1979. "Weakness of Will, Commensurability, and the Objects of Deliberation and Desire." *Proceedings of the Aristotelian Society* 79 (8):251-277.
- Williams, Bernard. 1973a. "The Makropulos Case: Reflections on the Tedium of Immortality." In *Problems of the Self*. Cambridge: Cambridge University Press.
- Williams, Bernard. 1973b. "Morality and the Emotions." In *Problems of the Self*. Cambridge: Cambridge University Press.
- Williams, Bernard. 2005. *Philosophy as a Humanistic Discipline*. Princeton: Princeton University Press.
- Wills, Gary. 2002. "Augustine's Pears and the Nature of Sin." *Arion* 10 (1).
- Wilson, John. 1988. "Why Forgiveness Requires Repentance." *Philosophy* 63.
- Wolf, Susan. 1982. "Moral Saints." *The Journal of Philosophy* 79 (8):419-439.
- Wolf, Susan. 1986. "Above and Below the Line of Duty." *Philosophical Topics* XIV (2).
- Zimmerman, Michael. 2011. "Partiality and Intrinsic Value." *Mind* 120 (478).